

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
16 January 2003 (16.01.2003)

PCT

(10) International Publication Number
WO 03/004523 A1

(51) International Patent Classification⁷: **C07K 14/435**,
C12N 15/52, 5/10, 9/00, C12Q 1/68, G01N 33/53, 33/573,
A61P 9/10

(21) International Application Number: PCT/EP02/07156

(22) International Filing Date: 28 June 2002 (28.06.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/301,841 2 July 2001 (02.07.2001) US
60/338,651 11 December 2001 (11.12.2001) US
60/375,014 25 April 2002 (25.04.2002) US

(71) Applicant (for all designated States except US): **BAYER AKTIENGESELLSCHAFT** [DE/DE]; 51368 Leverkusen (DE).

(72) Inventor; and

(75) Inventor/Applicant (for US only): **ZHU, Zhimin** [CN/US]; 45 Hinckley Road, Waban, MA 02468 (US).

(74) Common Representative: **BAYER AKTIENGESELLSCHAFT**; 51368 Leverkusen (DE).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for all designations

Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: REGULATION OF HUMAN CITRON RHO/RAC-INTERACTING KINASE

(57) Abstract: Reagents that regulate human CR1K and reagents which bind to human CR1K gene products can play a role in preventing, ameliorating, or correcting dysfunctions or diseases including, but not limited to, obesity, a CNS disorder or COPD.

WO 03/004523 A1

REGULATION OF HUMAN CITRON RHO/RAC-INTERACTING KINASE

This application incorporates by reference and claims the benefit of co-pending provisional applications Serial No. 60/301,841 filed July 2, 2001, Serial No.
5 60/338,651 filed December 11, 2001 and Serial No. 60/375,014 filed April 25, 2002

TECHNICAL FIELD OF THE INVENTION

The invention relates to the regulation of human citron rho/rac-interacting kinase
10 (CRIK).

BACKGROUND OF THE INVENTION

Kinases are involved in a variety of disease processes. There is a need in the art to
15 identify related enzymes, which can be regulated for therapeutic effects.

SUMMARY OF THE INVENTION

It is an object of the invention to provide reagents and methods of regulating a
20 human CRIK. This and other objects of the invention are provided by one or more of the embodiments described below.

One embodiment of the invention is a human citron rho/rac-interacting kinase polypeptide comprising an amino acid sequence selected from the group consisting
25 of:

amino acid sequences which are at least about 97% identical to the amino acid sequence shown in SEQ ID NO: 2; and

30 the amino acid sequence shown in SEQ ID NO: 2.

- 2 -

Yet another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a human citron rho/rac-interacting kinase polypeptide comprising an amino acid sequence selected from the group consisting of:

5

amino acid sequences which are at least about 97% identical to the amino acid sequence shown in SEQ ID NO: 2; and

the amino acid sequence shown in SEQ ID NO: 2.

10

Binding between the test compound and the human citron rho/rac-interacting kinase polypeptide is detected. A test compound which binds to the human citron rho/rac-interacting kinase polypeptide is thereby identified as a potential agent for decreasing extracellular matrix degradation. The agent can work by decreasing the activity of the human citron rho/rac-interacting kinase.

15

Another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a polynucleotide encoding a human citron rho/rac-interacting kinase polypeptide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

20

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1;

25

the nucleotide sequence shown in SEQ ID NO: 1;

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 24; and

30

the nucleotide sequence shown in SEQ ID NO: 24.

Binding of the test compound to the polynucleotide is detected. A test compound which binds to the polynucleotide is identified as a potential agent for decreasing extracellular matrix degradation. The agent can work by decreasing the amount of the human citron rho/rac-interacting kinase through interacting with the human citron rho/rac-interacting kinase mRNA.

Another embodiment of the invention is a method of screening for agents which regulate extracellular matrix degradation. A test compound is contacted with a human citron rho/rac-interacting kinase polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 97% identical to the amino acid sequence shown in SEQ ID NO: 2; and

the amino acid sequence shown in SEQ ID NO: 2.

A human citron rho/rac-interacting kinase activity of the polypeptide is detected. A test compound which increases human citron rho/rac-interacting kinase activity of the polypeptide relative to human citron rho/rac-interacting kinase activity in the absence of the test compound is thereby identified as a potential agent for increasing extracellular matrix degradation. A test compound which decreases human citron rho/rac-interacting kinase activity of the polypeptide relative to human citron rho/rac-interacting kinase activity in the absence of the test compound is thereby identified as a potential agent for decreasing extracellular matrix degradation.

Even another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a human citron rho/rac-interacting kinase product of a polynucleotide which comprises a nucleotide sequence selected from the group consisting of:

- 4 -

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1;

the nucleotide sequence shown in SEQ ID NO: 1;

5

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 24; and

the nucleotide sequence shown in SEQ ID NO: 24.

10

Binding of the test compound to the human citron rho/rac-interacting kinase product is detected. A test compound which binds to the human citron rho/rac-interacting kinase product is thereby identified as a potential agent for decreasing extracellular matrix degradation.

15

Still another embodiment of the invention is a method of reducing extracellular matrix degradation. A cell is contacted with a reagent which specifically binds to a polynucleotide encoding a human citron rho/rac-interacting kinase polypeptide or the product encoded by the polynucleotide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

20

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1;

25

the nucleotide sequence shown in SEQ ID NO: 1;

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 24; and

30

the nucleotide sequence shown in SEQ ID NO: 24.

Human citron rho/rac-interacting kinase activity in the cell is thereby decreased.

The invention thus provides a human CRIK that can be used to identify test compounds that may act, for example, as activators or inhibitors at the enzyme's active site. Human CRIK and fragments thereof also are useful in raising specific antibodies that can block the enzyme and effectively reduce its activity.

BRIEF DESCRIPTION OF THE DRAWINGS

- 10 Fig. 1 shows the DNA-sequence encoding a citron rho/rac-interacting kinase Polypeptide (SEQ ID NO:1).
- Fig. 2 shows the amino acid sequence deduced from the DNA-sequence of Fig.1 (SEQ ID NO:2).
- Fig. 3 shows the amino acid sequence of the protein identified by trembl|AF086824|AF086824_1 (SEQ ID NO:3).
- 15 Fig. 4 shows the DNA-sequence encoding a citron rho/rac-interacting kinase Polypeptide (SEQ ID NO:4).
- Fig. 5 shows the amino acid sequence of the protein identified by swiss|O14578|CTRO_HUMAN (SEQ ID NO:5).
- 20 Fig. 6 shows the DNA-sequence of a protein identified by trembl|AB023166|AB023166_1 (SEQ ID NO:6).
- Fig. 7 shows the amino acid sequence of the protein identified by swissnew|P54265|DMK_MOUSE (SEQ ID NO:7).
- Fig. 8 shows the BLASTP - alignment of 543_Protein (SEQ ID NO:2) against trembl|AF086824|AF086824_1 (SEQ ID NO:3).
- 25 Fig. 9 shows the BLASTP - alignment of 543_Protein (SEQ ID NO:2) against swiss|O14578|CTRO_HUMAN (SEQ ID NO:5).
- 30 Fig. 10 shows the BLASTP - alignment of 543_Protein (SEQ ID NO:2) against aageneseq|AAB43359|AAB43359.

- 6 -

- Fig. 11 shows the BLASTP - alignment of 543_Protein (SEQ ID NO:2) against trembl|AB023166|AB023166_1 (SEQ ID NO:6).
- 5 Fig. 12 shows the BLASTP - alignment of 543_Protein (SEQ ID NO:2) against swissnew|P54265|DMK_MOUSE (SEQ ID NO:7).
- Fig. 13 shows the BLASTP - alignment of 543_Protein (SEQ ID NO:2) against pdb|1CDK|1CDK-A.
- 10 Fig. 14 shows the HMMPFAM - alignment of 543_Protein (SEQ ID NO:2) against pfam|hmm|pkinase.
- Fig. 15 shows the HMMPFAM - alignment of 543_Protein (SEQ ID NO:2) against pfam|hmm|PH.
- Fig. 16 shows the HMMPFAM - alignment of 543_Protein (SEQ ID NO:2) against pfam|hmm|CNH.
- 15 Fig. 17 shows the HMMPFAM - alignment of 543_Protein (SEQ ID NO:2) against pfam|hmm|DAG_PE-bind.
- Fig. 18 shows the HMMPFAM - alignment of 543_Protein (SEQ ID NO:2) against pfam|hmm|pkinase_C.
- Fig. 19 shows the Prosite search results.
- 20 Fig. 20 shows the Genewise output.
- Fig. 21 shows the Relative expression of human citron rho/rac-interacting kinase.
- Fig. 22 shows the TBLASTN - alignment of 543_Protein against emnew|AX166510|AX166510 Sequence 1 from Patent WO0138503.//:gbnew|AX166510|AX166510 Sequence Patent WO0138503.
- 25 Fig. 23 shows the TBLASTN - alignment of 543_Protein against BAYER_LIB_DNA|wu_37300600 Bayer Corp Pharma Proprietary OP Library: Fat Rat Hypothalamus
- 30 Linda Oct 15 15:45:51 EDT 1999

Fig. 24 shows the DNA-sequence encoding a citron rho/rac-interacting kinase Polypeptide.

DETAILED DESCRIPTION OF THE INVENTION

5

The invention relates to an isolated polynucleotide from the group consisting of:

- a) a polynucleotide encoding a human citron rho/rac-interacting kinase polypeptide comprising an amino acid sequence selected from the group consisting of:
10 amino acid sequences which are at least about 97% identical to the amino acid sequence shown in SEQ ID NO: 2; and the amino acid sequence shown in SEQ ID NO: 2.
- b) a polynucleotide comprising the sequence of SEQ ID NOS: 1 or 24;
- c) a polynucleotide which hybridizes under stringent conditions to a
15 polynucleotide specified in (a) and (b) and encodes a human citron rho/rac-interacting kinase polypeptide;
- d) a polynucleotide the sequence of which deviates from the polynucleotide sequences specified in (a) to (c) due to the degeneration of the genetic code and encodes a human citron rho/rac-
20 interacting kinase polypeptide; and
- e) a polynucleotide which represents a fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (d) and encodes a human citron rho/rac-interacting kinase polypeptide.

25 Furthermore, it has been discovered by the present applicant that a novel CRIK, particularly a human CRIK, can be used in therapeutic methods to treat obesity, a CNS disorder, diabetes or COPD. Human CRIK comprises the amino acid sequence shown in SEQ ID NO:2. A coding sequence for human CRIK is shown in SEQ ID NO:1. This sequence is contained within the longer sequence shown in SEQ ID
30 NO:4, which is located on chromosome 12q24.2. Related ESTs are expressed in bone marrow, denis_drash (pediatric kidney tumors), epithelioid carcinoma

(pancreas), colon_ins (colon cancer cell line), uterus_tumor, glioblastoma with EGFR amplification, colon, nervous, nervous tumor, and bladder_tumor.

Human CRIK is 96% identical over 2056 amino acids to
5 trembl|AF086824|AF086824_1 (SEQ ID NO:3) (FIG. 1), 100% identical over 1286
amino acids to swiss|O14578|CTRO_HUMAN (SEQ ID NO:5) (FIG. 2), 100%
identical over 1286 amino acids to SEQ ID NO:6246 of
aageneseq|AAB43359|AAB43359 (FIG. 3), 100% over 940 amino acids to
trembl|AB023166|AB023166_1 (SEQ ID NO:6) (FIG. 4), and 38% identical over
10 522 amino acids to swissnew|P54265|DMK_MOUSE (SEQ ID NO:7) (FIG. 5).

Human CRIK of the invention is expected to be useful for the same purposes as
previously identified CRIK enzymes. Human CRIK is believed to be useful in
therapeutic methods to treat disorders such as CNS disorders, obesity, and COPD.
15 Human CRIK also can be used to screen for human CRIK activators and inhibitors.

Polypeptides

Human CRIK polypeptides according to the invention comprise at least 6, 10, 15, 20,
20 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450,
475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 725, 750, 775, 800, 825, 850, 875,
900, 925, 950, 975, 1000, 1025, 1050, 1075, 1100, 1125, 1150, 1175, 1200, 1225,
1250, 1275, 1300, 1325, 1350, 1375, 1400, 1425, 1450, 1475, 1500, 1525, 1550,
1575, 1600, 1625, 1650, 1675, 1700, 1725, 1750, 1775, 1800, 1825, 1850, 1875,
25 1900, 1925, 1950, 1975, 2000, 2025, 2050, or 2054 contiguous amino acids selected
from the amino acid sequence shown in SEQ ID NO:2 or a biologically active variant
thereof, as defined below. A CRIK polypeptide of the invention therefore can be a
portion of a CRIK protein, a full-length CRIK protein, or a fusion protein comprising
all or a portion of a CRIK protein.

Biologically Active Variants

Human CRIK polypeptide variants which are biologically active, *e.g.*, retain enzymatic activity, also are human CRIK polypeptides. Preferably, naturally or
5 non-naturally occurring human CRIK polypeptide variants have amino acid sequences which are at least about 97, 98, or 99% identical to the amino acid sequence shown in SEQ ID NO:2 or a fragment thereof. Percent identity between a putative human CRIK polypeptide variant and an amino acid sequence of SEQ ID NO:2 is determined by conventional methods. See, for example, Altschul *et al.*, *Bull.*
10 *Math. Bio.* 48:603 (1986), and Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1992). Briefly, two amino acid sequences are aligned to optimize the alignment scores using a gap opening penalty of 10, a gap extension penalty of 1, and the "BLOSUM62" scoring matrix of Henikoff & Henikoff, 1992.

15 Those skilled in the art appreciate that there are many established algorithms available to align two amino acid sequences. The "FASTA" similarity search algorithm of Pearson & Lipman is a suitable protein alignment method for examining the level of identity shared by an amino acid sequence disclosed herein and the amino acid sequence of a putative variant. The FASTA algorithm is described by
20 Pearson & Lipman, *Proc. Nat'l Acad. Sci. USA* 85:2444(1988), and by Pearson, *Meth. Enzymol.* 183:63 (1990). Briefly, FASTA first characterizes sequence similarity by identifying regions shared by the query sequence (*e.g.*, SEQ ID NO: 2) and a test sequence that have either the highest density of identities (if the *ktup* variable is 1) or pairs of identities (if *ktup*=2), without considering conservative
25 amino acid substitutions, insertions, or deletions. The ten regions with the highest density of identities are then rescored by comparing the similarity of all paired amino acids using an amino acid substitution matrix, and the ends of the regions are "trimmed" to include only those residues that contribute to the highest score. If there are several regions with scores greater than the "cutoff" value (calculated by a
30 predetermined formula based upon the length of the sequence the *ktup* value), then the trimmed initial regions are examined to determine whether the regions can be

joined to form an approximate alignment with gaps. Finally, the highest scoring regions of the two amino acid sequences are aligned using a modification of the Needleman-Wunsch- Sellers algorithm (Needleman & Wunsch, *J. Mol. Biol.* 48:444 (1970); Sellers, *SLAM J. Appl. Math.* 26:787 (1974)), which allows for amino acid
5 insertions and deletions. Preferred parameters for FASTA analysis are: ktup=1, gap opening penalty=10, gap extension penalty=1, and substitution matrix=BLOSUM62. These parameters can be introduced into a FASTA program by modifying the scoring matrix file ("SMATRIX"), as explained in Appendix 2 of Pearson, *Meth. Enzymol.* 183:63 (1990).

10 FASTA can also be used to determine the sequence identity of nucleic acid molecules using a ratio as disclosed above. For nucleotide sequence comparisons, the ktup value can range between one to six, preferably from three to six, most preferably three, with other parameters set as default.

15 Variations in percent identity can be due, for example, to amino acid substitutions, insertions, or deletions. Amino acid substitutions are defined as one for one amino acid replacements. They are conservative in nature when the substituted amino acid has similar structural and/or chemical properties. Examples of conservative
20 replacements are substitution of a leucine with an isoleucine or valine, an aspartate with a glutamate, or a threonine with a serine.

Amino acid insertions or deletions are changes to or within an amino acid sequence. They typically fall in the range of about 1 to 5 amino acids. Guidance in determining
25 which amino acid residues can be substituted, inserted, or deleted without abolishing biological or immunological activity of a human CRIK polypeptide can be found using computer programs well known in the art, such as DNASTAR software.

30 The invention additionally, encompasses CRIK polypeptides that are differentially modified during or after translation, *e.g.*, by glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups,

proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. Any of numerous chemical modifications can be carried out by known techniques including, but not limited, to specific chemical cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease, NaBH₄, acetylation, formylation, oxidation, reduction, metabolic synthesis in the presence of tunicamycin, etc.

Additional post-translational modifications encompassed by the invention include, for example, *e.g.*, N-linked or O-linked carbohydrate chains, processing of N-terminal or C-terminal ends), attachment of chemical moieties to the amino acid backbone, chemical modifications of N-linked or O-linked carbohydrate chains, and addition or deletion of an N-terminal methionine residue as a result of prokaryotic host cell expression. The CRIK polypeptides may also be modified with a detectable label, such as an enzymatic, fluorescent, isotopic or affinity label to allow for detection and isolation of the protein.

The invention also provides chemically modified derivatives of CRIK polypeptides that may provide additional advantages such as increased solubility, stability and circulating time of the polypeptide, or decreased immunogenicity (see U.S. Patent No. 4,179,337). The chemical moieties for derivitization can be selected from water soluble polymers such as polyethylene glycol, ethylene glycol/propylene glycol copolymers, carboxymethylcellulose, dextran, polyvinyl alcohol, and the like. The polypeptides can be modified at random or predetermined positions within the molecule and can include one, two, three, or more attached chemical moieties.

Whether an amino acid change or a polypeptide modification results in a biologically active CRIK polypeptide can readily be determined by assaying for enzymatic activity, as described for example, in Di Cunto F. *et al.*, J Biol Chem. 1998 Nov 6;273(45):29706-11.

Fusion Proteins

Fusion proteins are useful for generating antibodies against CRIK polypeptide amino acid sequences and for use in various assay systems. For example, fusion proteins can be used to identify proteins that interact with portions of a CRIK polypeptide. Protein affinity chromatography or library-based assays for protein-protein interactions, such as the yeast two-hybrid or phage display systems, can be used for this purpose. Such methods are well known in the art and also can be used as drug screens.

A CRIK polypeptide fusion protein comprises two polypeptide segments fused together by means of a peptide bond. The first polypeptide segment comprises at least 6, 10, 15, 20, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 725, 750, 775, 800, 825, 850, 875, 900, 925, 950, 975, 1000, 1025, 1050, 1075, 1100, 1125, 1150, 1175, 1200, 1225, 1250, 1275, 1300, 1325, 1350, 1375, 1400, 1425, 1450, 1475, 1500, 1525, 1550, 1575, 1600, 1625, 1650, 1675, 1700, 1725, 1750, 1775, 1800, 1825, 1850, 1875, 1900, 1925, 1950, 1975, 2000, 2025, 2050, or 2054 contiguous amino acids of SEQ ID NO:2 or of a biologically active variant, such as those described above. The first polypeptide segment also can comprise full-length CRIK protein.

The second polypeptide segment can be a full-length protein or a protein fragment. Proteins commonly used in fusion protein construction include β -galactosidase, β -glucuronidase, green fluorescent protein (GFP), autofluorescent proteins, including blue fluorescent protein (BFP), glutathione-S-transferase (GST), luciferase, horseradish peroxidase (HRP), and chloramphenicol acetyltransferase (CAT). Additionally, epitope tags are used in fusion protein constructions, including histidine (His) tags, FLAG tags, influenza hemagglutinin (HA) tags, Myc tags, VSV-G tags, and thioredoxin (Trx) tags. Other fusion constructions can include maltose binding protein (MBP), S-tag, Lex a DNA binding domain (DBD) fusions, GAL4

- 13 -

DNA binding domain fusions, and herpes simplex virus (HSV) BP16 protein fusions. A fusion protein also can be engineered to contain a cleavage site located between the CRIK polypeptide-encoding sequence and the heterologous protein sequence, so that the CRIK polypeptide can be cleaved and purified away from the heterologous moiety.

A fusion protein can be synthesized chemically, as is known in the art. Preferably, a fusion protein is produced by covalently linking two polypeptide segments or by standard procedures in the art of molecular biology. Recombinant DNA methods can be used to prepare fusion proteins, for example, by making a DNA construct which comprises coding sequences selected from SEQ ID NO:1 in proper reading frame with nucleotides encoding the second polypeptide segment and expressing the DNA construct in a host cell, as is known in the art. Many kits for constructing fusion proteins are available from companies such as Promega Corporation (Madison, WI), Stratagene (La Jolla, CA), CLONTECH (Mountain View, CA), Santa Cruz Biotechnology (Santa Cruz, CA), MBL International Corporation (MIC; Watertown, MA), and Quantum Biotechnologies (Montreal, Canada; 1-888-DNA-KITS).

Identification of Species Homologs

Species homologs of human CRIK polypeptide can be obtained using CRIK polypeptide polynucleotides (described below) to make suitable probes or primers for screening cDNA expression libraries from other species, such as mice, monkeys, or yeast, identifying cDNAs which encode homologs of CRIK polypeptide, and expressing the cDNAs as is known in the art.

Polynucleotides

5 A CRIK polynucleotide can be single- or double-stranded and comprises a coding sequence or the complement of a coding sequence for a CRIK polypeptide. A coding sequence for human CRIK is shown in SEQ ID NO:1.

10 Degenerate nucleotide sequences encoding human CRIK polypeptides, as well as homologous nucleotide sequences which are at least about 50, 55, 60, 65, 70, preferably about 75, 90, 96, 98, or 99% identical to the nucleotide sequence shown in SEQ ID NO:1 or its complement also are CRIK polynucleotides. Percent sequence identity between the sequences of two polynucleotides is determined using computer programs such as ALIGN which employ the FASTA algorithm, using an affine gap search with a gap open penalty of -12 and a gap extension penalty of -2. Complementary DNA (cDNA) molecules, species homologs, and variants of CRIK polynucleotides that encode biologically active CRIK polypeptides also are CRIK polynucleotides. Polynucleotide fragments comprising at least 8, 9, 10, 11, 12, 15, 20, or 25 contiguous nucleotides of SEQ ID NO:1 or its complement also are CRIK polynucleotides. These fragments can be used, for example, as hybridization probes or as antisense oligonucleotides.

20

Identification of Polynucleotide Variants and Homologs

25 Variants and homologs of the CRIK polynucleotides described above also are CRIK polynucleotides. Typically, homologous CRIK polynucleotide sequences can be identified by hybridization of candidate polynucleotides to known CRIK polynucleotides under stringent conditions, as is known in the art. For example, using the following wash conditions--2X SSC (0.3 M NaCl, 0.03 M sodium citrate, pH 7.0), 0.1% SDS, room temperature twice, 30 minutes each; then 2X SSC, 0.1% SDS, 50 °C once, 30 minutes; then 2X SSC, room temperature twice, 10 minutes each--homologous sequences can be identified which contain at most about 25-30%

30

basepair mismatches. More preferably, homologous nucleic acid strands contain 15-25% basepair mismatches, even more preferably 5-15% basepair mismatches.

Species homologs of the CRIK polynucleotides disclosed herein also can be identified by making suitable probes or primers and screening cDNA expression libraries from other species, such as mice, monkeys, or yeast. Human variants of CRIK polynucleotides can be identified, for example, by screening human cDNA expression libraries. It is well known that the T_m of a double-stranded DNA decreases by 1-1.5 °C with every 1% decrease in homology (Bonner *et al.*, *J. Mol. Biol.* 81, 123 (1973). Variants of human CRIK polynucleotides or CRIK polynucleotides of other species can therefore be identified by hybridizing a putative homologous CRIK polynucleotide with a polynucleotide having a nucleotide sequence of SEQ ID NO:1 or the complement thereof to form a test hybrid. The melting temperature of the test hybrid is compared with the melting temperature of a hybrid comprising polynucleotides having perfectly complementary nucleotide sequences, and the number or percent of basepair mismatches within the test hybrid is calculated.

Nucleotide sequences which hybridize to CRIK polynucleotides or their complements following stringent hybridization and/or wash conditions also are CRIK polynucleotides. Stringent wash conditions are well known and understood in the art and are disclosed, for example, in Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, 2d ed., 1989, at pages 9.50-9.51.

Typically, for stringent hybridization conditions a combination of temperature and salt concentration should be chosen that is approximately 12-20 °C below the calculated T_m of the hybrid under study. The T_m of a hybrid between a CRIK polynucleotide having a nucleotide sequence shown in SEQ ID NO:1 or the complement thereof and a polynucleotide sequence which is at least about 50, preferably about 75, 90, 96, or 98% identical to one of those nucleotide sequences

- 16 -

can be calculated, for example, using the equation of Bolton and McCarthy, *Proc. Natl. Acad. Sci. U.S.A.* 48, 1390 (1962):

$$T_m = 81.5\text{ }^{\circ}\text{C} - 16.6(\log_{10}[\text{Na}^+]) + 0.41(\%G + C) - 0.63(\%\text{formamide}) - 600/l,$$

5 where l = the length of the hybrid in basepairs.

Stringent wash conditions include, for example, 4X SSC at 65 °C, or 50% formamide, 4X SSC at 42 °C, or 0.5X SSC, 0.1% SDS at 65 °C. Highly stringent wash conditions include, for example, 0.2X SSC at 65 °C.

10

Preparation of Polynucleotides

A CRIK polynucleotide can be isolated free of other cellular components such as membrane components, proteins, and lipids. Polynucleotides can be made by a cell and isolated using standard nucleic acid purification techniques, or synthesized using an amplification technique, such as the polymerase chain reaction (PCR), or by using an automatic synthesizer. Methods for isolating polynucleotides are routine and are known in the art. Any such technique for obtaining a polynucleotide can be used to obtain isolated CRIK polynucleotides. For example, restriction enzymes and probes can be used to isolate polynucleotide fragments, which comprise CRIK nucleotide sequences. Isolated polynucleotides are in preparations that are free or at least 70, 80, or 90% free of other molecules.

20

Human CRIK cDNA molecules can be made with standard molecular biology techniques, using CRIK mRNA as a template. Human CRIK cDNA molecules can thereafter be replicated using molecular biology techniques known in the art and disclosed in manuals such as Sambrook *et al.* (1989). An amplification technique, such as PCR, can be used to obtain additional copies of polynucleotides of the invention, using either human genomic DNA or cDNA as a template.

25

30

Alternatively, synthetic chemistry techniques can be used to synthesize CRIK polynucleotides. The degeneracy of the genetic code allows alternate nucleotide sequences to be synthesized which will encode a CRIK polypeptide having, for example, an amino acid sequence shown in SEQ ID NO:2 or a biologically active variant thereof.

Extending Polynucleotides

Various PCR-based methods can be used to extend the nucleic acid sequences disclosed herein to detect upstream sequences such as promoters and regulatory elements. For example, restriction-site PCR uses universal primers to retrieve unknown sequence adjacent to a known locus (Sarkar, *PCR Methods Applic.* 2, 318-322, 1993). Genomic DNA is first amplified in the presence of a primer to a linker sequence and a primer specific to the known region. The amplified sequences are then subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Inverse PCR also can be used to amplify or extend sequences using divergent primers based on a known region (Triglia *et al.*, *Nucleic Acids Res.* 16, 8186, 1988). Primers can be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences Inc., Plymouth, Minn.), to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68-72 °C. The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template.

Another method which can be used is capture PCR, which involves PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA (Lagerstrom *et al.*, *PCR Methods Applic.* 1, 111-119,

1991). In this method, multiple restriction enzyme digestions and ligations also can be used to place an engineered double-stranded sequence into an unknown fragment of the DNA molecule before performing PCR.

5 Another method which can be used to retrieve unknown sequences is that of Parker *et al.*, *Nucleic Acids Res.* 19, 3055-3060, 1991). Additionally, PCR, nested primers, and PROMOTERFINDER libraries (CLONTECH, Palo Alto, Calif.) can be used to walk genomic DNA (CLONTECH, Palo Alto, Calif.). This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

10

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. Randomly-primed libraries are preferable, in that they will contain more sequences which contain the 5' regions of genes. Use of a randomly primed library may be especially preferable for situations in which an
15 oligo d(T) library does not yield a full-length cDNA. Genomic libraries can be useful for extension of sequence into 5' non-transcribed regulatory regions.

20

Commercially available capillary electrophoresis systems can be used to analyze the size or confirm the nucleotide sequence of PCR or sequencing products. For
25 example, capillary sequencing can employ flowable polymers for electrophoretic separation, four different fluorescent dyes (one for each nucleotide) that are laser activated, and detection of the emitted wavelengths by a charge coupled device camera. Output/light intensity can be converted to electrical signal using appropriate software (*e.g.* GENOTYPER and Sequence NAVIGATOR, Perkin Elmer), and the
entire process from loading of samples to computer analysis and electronic data display can be computer controlled. Capillary electrophoresis is especially preferable for the sequencing of small pieces of DNA that might be present in limited amounts in a particular sample.

Obtaining Polypeptides

Human CRIK polypeptides can be obtained, for example, by purification from human cells, by expression of CRIK polynucleotides, or by direct chemical synthesis.

5

Protein Purification

Human CRIK polypeptides can be purified from any cell that expresses the polypeptide, including host cells that have been transfected with CRIK expression constructs. A purified CRIK polypeptide is separated from other compounds that normally associate with the CRIK polypeptide in the cell, such as certain proteins, carbohydrates, or lipids, using methods well-known in the art. Such methods include, but are not limited to, size exclusion chromatography, ammonium sulfate fractionation, ion exchange chromatography, affinity chromatography, and preparative gel electrophoresis. A preparation of purified CRIK polypeptides is at least 80% pure; preferably, the preparations are 90%, 95%, or 99% pure. Purity of the preparations can be assessed by any means known in the art, such as SDS-polyacrylamide gel electrophoresis.

10

15

Expression of Polynucleotides

20

To express a CRIK polynucleotide, the polynucleotide can be inserted into an expression vector that contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods that are well known to those skilled in the art can be used to construct expression vectors containing sequences encoding CRIK polypeptides and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. Such techniques are described, for example, in Sambrook *et al.* (1989) and in Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, N.Y., 1989.

25

30

- 20 -

- A variety of expression vector/host systems can be utilized to contain and express sequences encoding a CRIK polypeptide. These include, but are not limited to, microorganisms, such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression
5 vectors, insect cell systems infected with virus expression vectors (*e.g.*, baculovirus), plant cell systems transformed with virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (*e.g.*, Ti or pBR322 plasmids), or animal cell systems.
- 10 The control elements or regulatory sequences are those non-translated regions of the vector -- enhancers, promoters, 5' and 3' untranslated regions -- which interact with host cellular proteins to carry out transcription and translation. Such elements can vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including
15 constitutive and inducible promoters, can be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the BLUESCRIPT phagemid (Stratagene, LaJolla, Calif.) or pSPORT1 plasmid (Life Technologies) and the like can be used. The baculovirus polyhedrin promoter can be used in insect cells. Promoters or enhancers derived from the genomes of plant
20 cells (*e.g.*, heat shock, RUBISCO, and storage protein genes) or from plant viruses (*e.g.*, viral promoters or leader sequences) can be cloned into the vector. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are preferable. If it is necessary to generate a cell line that contains multiple copies of a nucleotide sequence encoding a CRIK polypeptide, vectors based on
25 SV40 or EBV can be used with an appropriate selectable marker.

Bacterial and Yeast Expression Systems

- 30 In bacterial systems, a number of expression vectors can be selected depending upon the use intended for the CRIK polypeptide. For example, when a large quantity of a CRIK polypeptide is needed for the induction of antibodies, vectors which direct

high level expression of fusion proteins that are readily purified can be used. Such vectors include, but are not limited to, multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene). In a BLUESCRIPT vector, a sequence encoding the CRIK polypeptide can be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of β -galactosidase so that a hybrid protein is produced. pIN vectors (Van Heeke & Schuster, *J. Biol. Chem.* 264, 5503-5509, 1989) or pGEX vectors (Promega, Madison, Wis.) also can be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems can be designed to include heparin, thrombin, or factor Xa protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH can be used. For reviews, see Ausubel *et al.* (1989) and Grant *et al.*, *Methods Enzymol.* 153, 516-544, 1987.

20

Plant and Insect Expression Systems

If plant expression vectors are used, the expression of sequences encoding CRIK polypeptides can be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV can be used alone or in combination with the omega leader sequence from TMV (Takamatsu, *EMBO J.* 6, 307-311, 1987). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters can be used (Coruzzi *et al.*, *EMBO J.* 3, 1671-1680, 1984; Broglie *et al.*, *Science* 224, 838-843, 1984; Winter *et al.*, *Results Probl. Cell Differ.* 17, 85-105, 1991). These constructs can be introduced into plant cells by direct DNA transformation or by pathogen-mediated transfection. Such

techniques are described in a number of generally available reviews (e.g., Hobbs or Murray, in MCGRAW HILL YEARBOOK OF SCIENCE AND TECHNOLOGY, McGraw Hill, New York, N.Y., pp. 191-196, 1992).

5 An insect system also can be used to express a CRIK polypeptide. For example, in one such system *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in *Trichoplusia* larvae. Sequences encoding CRIK polypeptides can be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the
10 polyhedrin promoter. Successful insertion of CRIK polypeptides will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses can then be used to infect *S. frugiperda* cells or *Trichoplusia* larvae in which CRIK polypeptides can be expressed (Engelhard *et al.*, *Proc. Nat. Acad. Sci.* 91, 3224-3227, 1994).

15

Mammalian Expression Systems

A number of viral-based expression systems can be used to express CRIK polypeptides in mammalian host cells. For example, if an adenovirus is used as an
20 expression vector, sequences encoding CRIK polypeptides can be ligated into an adenovirus transcription/translation complex comprising the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome can be used to obtain a viable virus that is capable of expressing a CRIK polypeptide in infected host cells (Logan & Shenk, *Proc. Natl. Acad. Sci.* 81, 3655-3659, 1984). If desired, transcription enhancers, such as the Rous sarcoma
25 virus (RSV) enhancer, can be used to increase expression in mammalian host cells.

Human artificial chromosomes (HACs) also can be used to deliver larger fragments of DNA than can be contained and expressed in a plasmid. HACs of 6M to 10M are
30 constructed and delivered to cells via conventional delivery methods (e.g., liposomes, polycationic amino polymers, or vesicles).

Specific initiation signals also can be used to achieve more efficient translation of sequences encoding CRIK polypeptides. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding a CRIK polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals (including the ATG initiation codon) should be provided. The initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons can be of various origins, both natural and synthetic. The efficiency of expression can be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used (see Scharf *et al.*, *Results Probl. Cell Differ.* 20, 125-162, 1994).

15

Host Cells

A host cell strain can be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed CRIK polypeptide in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the polypeptide also can be used to facilitate correct insertion, folding and/or function. Different host cells that have specific cellular machinery and characteristic mechanisms for post-translational activities (*e.g.*, CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC; 10801 University Boulevard, Manassas, VA 20110-2209) and can be chosen to ensure the correct modification and processing of the foreign protein.

30

Stable expression is preferred for long-term, high-yield production of recombinant proteins. For example, cell lines which stably express CRIK polypeptides can be

transformed using expression vectors which can contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells can be allowed to grow for 1-2 days in an enriched medium before they are switched to a selective medium. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced CRIK sequences. Resistant clones of stably transformed cells can be proliferated using tissue culture techniques appropriate to the cell type. See, for example, ANIMAL CELL CULTURE, R.I. Freshney, ed., 1986.

Any number of selection systems can be used to recover transformed cell lines.

These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler *et al.*, *Cell* 11, 223-32, 1977) and adenine phosphoribosyltransferase (Lowy *et al.*, *Cell* 22, 817-23, 1980) genes which can be employed in *tk* or *aprt* cells, respectively. Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate (Wigler *et al.*, *Proc. Natl. Acad. Sci.* 77, 3567-70, 1980), *npt* confers resistance to the aminoglycosides, neomycin and G-418 (Colbere-Garapin *et al.*, *J. Mol. Biol.* 150, 1-14, 1981), and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murray, 1992, *supra*). Additional selectable genes have been described. For example, *trpB* allows cells to utilize indole in place of tryptophan, or *hisD*, which allows cells to utilize histinol in place of histidine (Hartman & Mulligan, *Proc. Natl. Acad. Sci.* 85, 8047-51, 1988). Visible markers such as anthocyanins, β -glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, can be used to identify transformants and to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes *et al.*, *Methods Mol. Biol.* 55, 121-131, 1995).

Detecting Expression

Although the presence of marker gene expression suggests that the CRIK polynucleotide is also present, its presence and expression may need to be confirmed. For example, if a sequence encoding a CRIK polypeptide is inserted within a marker gene sequence, transformed cells containing sequences that encode a CRIK polypeptide can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding a CRIK polypeptide under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the CRIK polynucleotide.

Alternatively, host cells which contain a CRIK polynucleotide and which express a CRIK polypeptide can be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques that include membrane, solution, or chip-based technologies for the detection and/or quantification of nucleic acid or protein. For example, the presence of a polynucleotide sequence encoding a CRIK polypeptide can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes or fragments or fragments of polynucleotides encoding a CRIK polypeptide. Nucleic acid amplification-based assays involve the use of oligonucleotides selected from sequences encoding a CRIK polypeptide to detect transformants that contain a CRIK polynucleotide.

A variety of protocols for detecting and measuring the expression of a CRIK polypeptide, using either polyclonal or monoclonal antibodies specific for the polypeptide, are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay using monoclonal antibodies reactive to two non-interfering epitopes on a CRIK polypeptide can be used, or a competitive binding assay can be employed. These and other assays are described in Hampton *et*

- 26 -

al., SEROLOGICAL METHODS: A LABORATORY MANUAL, APS Press, St. Paul, Minn., 1990) and Maddox *et al.*, *J. Exp. Med.* 158, 1211-1216, 1983).

5 A wide variety of labels and conjugation techniques are known by those skilled in the art and can be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding CRIK polypeptides include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, sequences encoding a CRIK polypeptide can be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and can be used to synthesize RNA probes *in vitro* by addition of labeled nucleotides and an appropriate RNA polymerase such as T7, T3, or SP6. These procedures can be conducted using a variety of commercially available kits (Amersham Pharmacia Biotech, Promega, and US Biochemical). Suitable reporter molecules or labels 15 which can be used for ease of detection include radionuclides, enzymes, and fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, co-factors, inhibitors, magnetic particles, and the like.

Expression and Purification of Polypeptides

20 Host cells transformed with nucleotide sequences encoding a CRIK polypeptide can be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The polypeptide produced by a transformed cell can be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode CRIK polypeptides can be designed to contain signal sequences which direct secretion of soluble CRIK polypeptides through a prokaryotic or eukaryotic cell membrane or which direct the membrane insertion of membrane-bound CRIK polypeptide.

30

As discussed above, other constructions can be used to join a sequence encoding a CRIK polypeptide to a nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). Inclusion of cleavable linker sequences such as those specific for Factor Xa or enterokinase (Invitrogen, San Diego, CA) between the purification domain and the CRIK polypeptide also can be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing a CRIK polypeptide and 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification by IMAC (immobilized metal ion affinity chromatography, as described in Porath *et al.*, *Prot. Exp. Purif.* 3, 263-281, 1992), while the enterokinase cleavage site provides a means for purifying the CRIK polypeptide from the fusion protein. Vectors that contain fusion proteins are disclosed in Kroll *et al.*, *DNA Cell Biol.* 12, 441-453, 1993.

Chemical Synthesis

Sequences encoding a CRIK polypeptide can be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers *et al.*, *Nucl. Acids Res. Symp. Ser.* 215-223, 1980; Horn *et al.* *Nucl. Acids Res. Symp. Ser.* 225-232, 1980). Alternatively, a CRIK polypeptide itself can be produced using chemical methods to synthesize its amino acid sequence, such as by direct peptide synthesis using solid-phase techniques (Merrifield, *J. Am. Chem. Soc.* 85, 2149-2154, 1963; Roberge *et al.*, *Science* 269, 202-204, 1995). Protein synthesis can be performed using manual techniques or by automation. Automated synthesis can be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Optionally, fragments of CRIK polypeptides can be separately synthesized and combined using chemical methods to produce a full-length molecule.

The newly synthesized peptide can be substantially purified by preparative high performance liquid chromatography (*e.g.*, Creighton, PROTEINS: STRUCTURES AND MOLECULAR PRINCIPLES, WH Freeman and Co., New York, N.Y., 1983). The
5 composition of a synthetic CRIK polypeptide can be confirmed by amino acid analysis or sequencing (*e.g.*, the Edman degradation procedure; *see* Creighton, *supra*). Additionally, any portion of the amino acid sequence of the CRIK polypeptide can be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins to produce a variant polypeptide or a
10 fusion protein.

Production of Altered Polypeptides

As will be understood by those of skill in the art, it may be advantageous to produce
15 CRIK polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce an RNA transcript having desirable properties, such as a half-life that is longer than that of a transcript generated from the naturally occurring sequence.

20 The nucleotide sequences disclosed herein can be engineered using methods generally known in the art to alter CRIK polypeptide-encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the polypeptide or mRNA product. DNA shuffling
25 by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides can be used to engineer the nucleotide sequences. For example, site-directed mutagenesis can be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, introduce mutations, and so forth.

30

Antibodies

Any type of antibody known in the art can be generated to bind specifically to an epitope of a CRIK polypeptide. "Antibody" as used herein includes intact immunoglobulin molecules, as well as fragments thereof, such as Fab, F(ab')₂, and Fv, which are capable of binding an epitope of a CRIK polypeptide. Typically, at least 6, 8, 10, or 12 contiguous amino acids are required to form an epitope. However, epitopes which involve non-contiguous amino acids may require more, *e.g.*, at least 15, 25, or 50 amino acids.

An antibody which specifically binds to an epitope of a CRIK polypeptide can be used therapeutically, as well as in immunochemical assays, such as Western blots, ELISAs, radioimmunoassays, immunohistochemical assays, immunoprecipitations, or other immunochemical assays known in the art. Various immunoassays can be used to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays are well known in the art. Such immunoassays typically involve the measurement of complex formation between an immunogen and an antibody that specifically binds to the immunogen.

Typically, an antibody which specifically binds to a CRIK polypeptide provides a detection signal at least 5-, 10-, or 20-fold higher than a detection signal provided with other proteins when used in an immunochemical assay. Preferably, antibodies which specifically bind to CRIK polypeptides do not detect other proteins in immunochemical assays and can immunoprecipitate a CRIK polypeptide from solution.

Human CRIK polypeptides can be used to immunize a mammal, such as a mouse, rat, rabbit, guinea pig, monkey, or human, to produce polyclonal antibodies. If desired, a CRIK polypeptide can be conjugated to a carrier protein, such as bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin. Depending on the host species, various adjuvants can be used to increase the immunological response.

- 30 -

Such adjuvants include, but are not limited to, Freund's adjuvant, mineral gels (e.g., aluminum hydroxide), and surface active substances (e.g. lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol). Among adjuvants used in humans, BCG (*bacilli Calmette-Guerin*) and *Corynebacterium parvum* are especially useful.

Monoclonal antibodies that specifically bind to a CRIK polypeptide can be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These techniques include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler *et al.*, *Nature* 256, 495-497, 1985; Kozbor *et al.*, *J. Immunol. Methods* 81, 31-42, 1985; Cote *et al.*, *Proc. Natl. Acad. Sci.* 80, 2026-2030, 1983; Cole *et al.*, *Mol. Cell Biol.* 62, 109-120, 1984).

In addition, techniques developed for the production of "chimeric antibodies," the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used (Morrison *et al.*, *Proc. Natl. Acad. Sci.* 81, 6851-6855, 1984; Neuberger *et al.*, *Nature* 312, 604-608, 1984; Takeda *et al.*, *Nature* 314, 452-454, 1985). Monoclonal and other antibodies also can be "humanized" to prevent a patient from mounting an immune response against the antibody when it is used therapeutically. Such antibodies may be sufficiently similar in sequence to human antibodies to be used directly in therapy or may require alteration of a few key residues. Sequence differences between rodent antibodies and human sequences can be minimized by replacing residues which differ from those in the human sequences by site directed mutagenesis of individual residues or by grating of entire complementarity determining regions. Alternatively, humanized antibodies can be produced using recombinant methods, as described in GB2188638B. Antibodies that specifically bind to a CRIK polypeptide can contain antigen binding sites which are either partially or fully humanized, as disclosed in U.S. 5,565,332.

Alternatively, techniques described for the production of single chain antibodies can be adapted using methods known in the art to produce single chain antibodies that specifically bind to CRIK polypeptides. Antibodies with related specificity, but of distinct idiotypic composition, can be generated by chain shuffling from random combinatorial immunoglobulin libraries (Burton, *Proc. Natl. Acad. Sci.* 88, 11120-23, 1991).

Single-chain antibodies also can be constructed using a DNA amplification method, such as PCR, using hybridoma cDNA as a template (Thirion *et al.*, 1996, *Eur. J. Cancer Prev.* 5, 507-11). Single-chain antibodies can be mono- or bispecific, and can be bivalent or tetravalent. Construction of tetravalent, bispecific single-chain antibodies is taught, for example, in Coloma & Morrison, 1997, *Nat. Biotechnol.* 15, 159-63. Construction of bivalent, bispecific single-chain antibodies is taught in Mallender & Voss, 1994, *J. Biol. Chem.* 269, 199-206.

A nucleotide sequence encoding a single-chain antibody can be constructed using manual or automated nucleotide synthesis, cloned into an expression construct using standard recombinant DNA methods, and introduced into a cell to express the coding sequence, as described below. Alternatively, single-chain antibodies can be produced directly using, for example, filamentous phage technology (Verhaar *et al.*, 1995, *Int. J. Cancer* 61, 497-501; Nicholls *et al.*, 1993, *J. Immunol. Meth.* 165, 81-91).

Antibodies which specifically bind to CRIK polypeptides also can be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi *et al.*, *Proc. Natl. Acad. Sci.* 86, 3833-3837, 1989; Winter *et al.*, *Nature* 349, 293-299, 1991).

Other types of antibodies can be constructed and used therapeutically in methods of the invention. For example, chimeric antibodies can be constructed as disclosed in

WO 93/03151. Binding proteins which are derived from immunoglobulins and which are multivalent and multispecific, such as the "diabodies" described in WO 94/13804, also can be prepared.

5 Antibodies according to the invention can be purified by methods well known in the art. For example, antibodies can be affinity purified by passage over a column to which a CRIK polypeptide is bound. The bound antibodies can then be eluted from the column using a buffer with a high salt concentration.

10 Antisense Oligonucleotides

Antisense oligonucleotides are nucleotide sequences that are complementary to a specific DNA or RNA sequence. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form complexes
15 and block either transcription or translation. Preferably, an antisense oligonucleotide is at least 11 nucleotides in length, but can be at least 12, 15, 20, 25, 30, 35, 40, 45, or 50 or more nucleotides long. Longer sequences also can be used. Antisense oligonucleotide molecules can be provided in a DNA construct and introduced into a cell as described above to decrease the level of CRIK gene products in the cell.

20

Antisense oligonucleotides can be deoxyribonucleotides, ribonucleotides, or a combination of both. Oligonucleotides can be synthesized manually or by an automated synthesizer, by covalently linking the 5' end of one nucleotide with the 3' end of another nucleotide with non-phosphodiester internucleotide linkages such
25 alkylphosphonates, phosphorothioates, phosphorodithioates, alkylphosphonothioates, alkylphosphonates, phosphoramidates, phosphate esters, carbamates, acetamidate, carboxymethyl esters, carbonates, and phosphate triesters. See Brown, *Meth. Mol. Biol.* 20, 1-8, 1994; Sonveaux, *Meth. Mol. Biol.* 26, 1-72, 1994; Uhlmann *et al.*, *Chem. Rev.* 90, 543-583, 1990.

30

Modifications of CRIK gene expression can be obtained by designing antisense oligonucleotides that will form duplexes to the control, 5', or regulatory regions of the CRIK gene. Oligonucleotides derived from the transcription initiation site, *e.g.*, between positions -10 and +10 from the start site, are preferred. Similarly, inhibition
5 can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or chaperons. Therapeutic advances using triplex DNA have been described in the literature (*e.g.*, Gee *et al.*, in Huber & Carr, MOLECULAR AND IMMUNOLOGIC APPROACHES, Futura
10 Publishing Co., Mt. Kisco, N.Y., 1994). An antisense oligonucleotide also can be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Precise complementarity is not required for successful complex formation between
15 an antisense oligonucleotide and the complementary sequence of a CRIK polynucleotide. Antisense oligonucleotides which comprise, for example, 2, 3, 4, or 5 or more stretches of contiguous nucleotides which are precisely complementary to a CRIK polynucleotide, each separated by a stretch of contiguous nucleotides which are not complementary to adjacent CRIK nucleotides, can provide sufficient targeting
20 specificity for CRIK mRNA. Preferably, each stretch of complementary contiguous nucleotides is at least 4, 5, 6, 7, or 8 or more nucleotides in length. Non-complementary intervening sequences are preferably 1, 2, 3, or 4 nucleotides in length. One skilled in the art can easily use the calculated melting point of an antisense-sense pair to determine the degree of mismatching which will be tolerated
25 between a particular antisense oligonucleotide and a particular CRIK polynucleotide sequence.

Antisense oligonucleotides can be modified without affecting their ability to hybridize to a CRIK polynucleotide. These modifications can be internal or at one or
30 both ends of the antisense molecule. For example, internucleoside phosphate linkages can be modified by adding cholesteryl or diamine moieties with varying

numbers of carbon residues between the amino groups and terminal ribose. Modified bases and/or sugars, such as arabinose instead of ribose, or a 3', 5'-substituted oligonucleotide in which the 3' hydroxyl group or the 5' phosphate group are substituted, also can be employed in a modified antisense oligonucleotide. These modified oligonucleotides can be prepared by methods well known in the art. See, e.g., Agrawal *et al.*, *Trends Biotechnol.* 10, 152-158, 1992; Uhlmann *et al.*, *Chem. Rev.* 90, 543-584, 1990; Uhlmann *et al.*, *Tetrahedron. Lett.* 215, 3539-3542, 1987.

Ribozymes

Ribozymes are RNA molecules with catalytic activity. See, e.g., Cech, *Science* 236, 1532-1539; 1987; Cech, *Ann. Rev. Biochem.* 59, 543-568; 1990, Cech, *Curr. Opin. Struct. Biol.* 2, 605-609; 1992, Couture & Stinchcomb, *Trends Genet.* 12, 510-515, 1996. Ribozymes can be used to inhibit gene function by cleaving an RNA sequence, as is known in the art (e.g., Haseloff *et al.*, U.S. Patent 5,641,673). The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Examples include engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of specific nucleotide sequences.

The coding sequence of a CRIK polynucleotide can be used to generate ribozymes that will specifically bind to mRNA transcribed from the CRIK polynucleotide. Methods of designing and constructing ribozymes which can cleave other RNA molecules in trans in a highly sequence specific manner have been developed and described in the art (see Haseloff *et al.* *Nature* 334, 585-591, 1988). For example, the cleavage activity of ribozymes can be targeted to specific RNAs by engineering a discrete "hybridization" region into the ribozyme. The hybridization region contains a sequence complementary to the target RNA and thus specifically hybridizes with the target (see, for example, Gerlach *et al.*, EP 321,201).

Specific ribozyme cleavage sites within a CRIK RNA target can be identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target RNA containing the cleavage site can be evaluated for secondary structural features which may render the target inoperable. Suitability of candidate CRIK RNA targets also can be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays. Longer complementary sequences can be used to increase the affinity of the hybridization sequence for the target. The hybridizing and cleavage regions of the ribozyme can be integrally related such that upon hybridizing to the target RNA through the complementary regions, the catalytic region of the ribozyme can cleave the target.

Ribozymes can be introduced into cells as part of a DNA construct. Mechanical methods, such as microinjection, liposome-mediated transfection, electroporation, or calcium phosphate precipitation, can be used to introduce a ribozyme-containing DNA construct into cells in which it is desired to decrease CRIK expression. Alternatively, if it is desired that the cells stably retain the DNA construct, the construct can be supplied on a plasmid and maintained as a separate element or integrated into the genome of the cells, as is known in the art. A ribozyme-encoding DNA construct can include transcriptional regulatory elements, such as a promoter element, an enhancer or UAS element, and a transcriptional terminator signal, for controlling transcription of ribozymes in the cells.

As taught in Haseloff *et al.*, U.S. Patent 5,641,673, ribozymes can be engineered so that ribozyme expression will occur in response to factors that induce expression of a target gene. Ribozymes also can be engineered to provide an additional level of regulation, so that destruction of mRNA occurs only when both a ribozyme and a target gene are induced in the cells.

Differentially Expressed Genes

Described herein are methods for the identification of genes whose products interact with human CRIK. Such genes may represent genes that are differentially expressed in disorders including, but not limited to, obesity, CNS disorders, and COPD. Further, such genes may represent genes that are differentially regulated in response to manipulations relevant to the progression or treatment of such diseases. Additionally, such genes may have a temporally modulated expression, increased or decreased at different stages of tissue or organism development. A differentially expressed gene may also have its expression modulated under control versus experimental conditions. In addition, the human CRIK gene or gene product may itself be tested for differential expression.

The degree to which expression differs in a normal versus a diseased state need only be large enough to be visualized via standard characterization techniques such as differential display techniques. Other such standard characterization techniques by which expression differences may be visualized include but are not limited to, quantitative RT (reverse transcriptase), PCR, and Northern analysis.

Identification of Differentially Expressed Genes

To identify differentially expressed genes total RNA or, preferably, mRNA is isolated from tissues of interest. For example, RNA samples are obtained from tissues of experimental subjects and from corresponding tissues of control subjects. Any RNA isolation technique that does not select against the isolation of mRNA may be utilized for the purification of such RNA samples. See, for example, Ausubel *et al.*, ed., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, Inc. New York, 1987-1993. Large numbers of tissue samples may readily be processed using techniques well known to those of skill in the art, such as, for example, the single-step RNA isolation process of Chomczynski, U.S. Patent 4,843,155.

Transcripts within the collected RNA samples that represent RNA produced by differentially expressed genes are identified by methods well known to those of skill in the art. They include, for example, differential screening (Tedder *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 85, 208-12, 1988), subtractive hybridization (Hedrick *et al.*,
5 *Nature* 308, 149-53; Lee *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 88, 2825, 1984), and, preferably, differential display (Liang & Pardee, *Science* 257, 967-71, 1992; U.S. Patent 5,262,311).

The differential expression information may itself suggest relevant methods for the
10 treatment of disorders involving the human CRIK. For example, treatment may include a modulation of expression of the differentially expressed genes and/or the gene encoding the human CRIK. The differential expression information may indicate whether the expression or activity of the differentially expressed gene or gene product or the human CRIK gene or gene product are up-regulated or down-
15 regulated.

Screening Methods

The invention provides assays for screening test compounds that bind to or modulate
20 the activity of a CRIK polypeptide or a CRIK polynucleotide. A test compound preferably binds to a CRIK polypeptide or polynucleotide. More preferably, a test compound decreases or increases enzymatic activity by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the test compound.

Test Compounds

Test compounds can be pharmacologic agents already known in the art or can be compounds previously unknown to have any pharmacological activity. The
30 compounds can be naturally occurring or designed in the laboratory. They can be isolated from microorganisms, animals, or plants, and can be produced

recombinantly, or synthesized by chemical methods known in the art. If desired, test compounds can be obtained using any of the numerous combinatorial library methods known in the art, including but not limited to, biological libraries, spatially addressable parallel solid phase or solution phase libraries, synthetic library methods requiring deconvolution, the "one-bead one-compound" library method, and synthetic library methods using affinity chromatography selection. The biological library approach is limited to polypeptide libraries, while the other four approaches are applicable to polypeptide, non-peptide oligomer, or small molecule libraries of compounds. See Lam, *Anticancer Drug Des.* 12, 145, 1997.

Methods for the synthesis of molecular libraries are well known in the art (see, for example, DeWitt *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 90, 6909, 1993; Erb *et al.* *Proc. Natl. Acad. Sci. U.S.A.* 91, 11422, 1994; Zuckermann *et al.*, *J. Med. Chem.* 37, 2678, 1994; Cho *et al.*, *Science* 261, 1303, 1993; Carell *et al.*, *Angew. Chem. Int. Ed. Engl.* 33, 2059, 1994; Carell *et al.*, *Angew. Chem. Int. Ed. Engl.* 33, 2061; Gallop *et al.*, *J. Med. Chem.* 37, 1233, 1994). Libraries of compounds can be presented in solution (see, e.g., Houghten, *BioTechniques* 13, 412-421, 1992), or on beads (Lam, *Nature* 354, 82-84, 1991), chips (Fodor, *Nature* 364, 555-556, 1993), bacteria or spores (Ladner, U.S. Patent 5,223,409), plasmids (Cull *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 89, 1865-1869, 1992), or phage (Scott & Smith, *Science* 249, 386-390, 1990; Devlin, *Science* 249, 404-406, 1990); Cwirla *et al.*, *Proc. Natl. Acad. Sci.* 97, 6378-6382, 1990; Felici, *J. Mol. Biol.* 222, 301-310, 1991; and Ladner, U.S. Patent 5,223,409).

High Throughput Screening

Test compounds can be screened for the ability to bind to CRIK polypeptides or polynucleotides or to affect CRIK activity or CRIK gene expression using high throughput screening. Using high throughput screening, many discrete compounds can be tested in parallel so that large numbers of test compounds can be quickly screened. The most widely established techniques utilize 96-well microtiter plates. The wells of the microtiter plates typically require assay volumes that range from 50

to 500 μ l. In addition to the plates, many instruments, materials, pipettors, robotics, plate washers, and plate readers are commercially available to fit the 96-well format.

Alternatively, "free format assays," or assays that have no physical barrier between
5 samples, can be used. For example, an assay using pigment cells (melanocytes) in a simple homogeneous assay for combinatorial peptide libraries is described by Jayawickreme *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 19, 1614-18 (1994). The cells are placed under agarose in petri dishes, then beads that carry combinatorial compounds are placed on the surface of the agarose. The combinatorial compounds are partially
10 released the compounds from the beads. Active compounds can be visualized as dark pigment areas because, as the compounds diffuse locally into the gel matrix, the active compounds cause the cells to change colors.

Another example of a free format assay is described by Chelsky, "Strategies for
15 Screening Combinatorial Libraries: Novel and Traditional Approaches," reported at the First Annual Conference of The Society for Biomolecular Screening in Philadelphia, Pa. (Nov. 7-10, 1995). Chelsky placed a simple homogenous enzyme assay for carbonic anhydrase inside an agarose gel such that the enzyme in the gel would cause a color change throughout the gel. Thereafter, beads carrying
20 combinatorial compounds via a photolinker were placed inside the gel and the compounds were partially released by UV-light. Compounds that inhibited the enzyme were observed as local zones of inhibition having less color change.

Yet another example is described by Salmon *et al.*, *Molecular Diversity* 2, 57-63
25 (1996). In this example, combinatorial libraries were screened for compounds that had cytotoxic effects on cancer cells growing in agar.

Another high throughput screening method is described in Beutel *et al.*, U.S. Patent 5,976,813. In this method, test samples are placed in a porous matrix. One or more
30 assay components are then placed within, on top of, or at the bottom of a matrix such as a gel, a plastic sheet, a filter, or other form of easily manipulated solid support.

When samples are introduced to the porous matrix they diffuse sufficiently slowly, such that the assays can be performed without the test samples running together.

Binding Assays

5

For binding assays, the test compound is preferably a small molecule that binds to and occupies, for example, the active site of the CRIK polypeptide, such that normal biological activity is prevented. Examples of such small molecules include, but are not limited to, small peptides or peptide-like molecules.

10

In binding assays, either the test compound or the CRIK polypeptide can comprise a detectable label, such as a fluorescent, radioisotopic, chemiluminescent, or enzymatic label, such as horseradish peroxidase, alkaline phosphatase, or luciferase. Detection of a test compound that is bound to the CRIK polypeptide can then be accomplished, for example, by direct counting of radioemission, by scintillation counting, or by determining conversion of an appropriate substrate to a detectable product.

15

Alternatively, binding of a test compound to a CRIK polypeptide can be determined without labeling either of the interactants. For example, a microphysiometer can be used to detect binding of a test compound with a CRIK polypeptide. A microphysiometer (*e.g.*, Cytosensor™) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a test compound and a CRIK polypeptide (McConnell *et al.*, *Science* 257, 1906-1912, 1992).

20

25

Determining the ability of a test compound to bind to a CRIK polypeptide also can be accomplished using a technology such as real-time Bimolecular Interaction Analysis (BIA) (Sjolander & Urbaniczky, *Anal. Chem.* 63, 2338-2345, 1991, and Szabo *et al.*, *Curr. Opin. Struct. Biol.* 5, 699-705, 1995). BIA is a technology for studying biospecific interactions in real time, without labeling any of the interactants

30

(e.g., BIAcore™). Changes in the optical phenomenon surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

5 In yet another aspect of the invention, a CRIK polypeptide can be used as a "bait protein" in a two-hybrid assay or three-hybrid assay (see, e.g., U.S. Patent 5,283,317; Zervos *et al.*, *Cell* 72, 223-232, 1993; Madura *et al.*, *J. Biol. Chem.* 268, 12046-12054, 1993; Bartel *et al.*, *BioTechniques* 14, 920-924, 1993; Iwabuchi *et al.*, *Oncogene* 8, 1693-1696, 1993; and Brent W094/10300), to identify other proteins
10 which bind to or interact with the CRIK polypeptide and modulate its activity.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. For example, in one construct, polynucleotide
15 encoding a CRIK polypeptide can be fused to a polynucleotide encoding the DNA binding domain of a known transcription factor (e.g., GAL-4). In the other construct a DNA sequence that encodes an unidentified protein ("prey" or "sample") can be fused to a polynucleotide that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact *in vivo*
20 to form an protein-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (e.g., LacZ), which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected, and cell colonies containing the functional
25 transcription factor can be isolated and used to obtain the DNA sequence encoding the protein that interacts with the CRIK polypeptide.

It may be desirable to immobilize either the CRIK polypeptide (or polynucleotide) or the test compound to facilitate separation of bound from unbound forms of one or
30 both of the interactants, as well as to accommodate automation of the assay. Thus, either the CRIK polypeptide (or polynucleotide) or the test compound can be bound

to a solid support. Suitable solid supports include, but are not limited to, glass or plastic slides, tissue culture plates, microtiter wells, tubes, silicon chips, or particles such as beads (including, but not limited to, latex, polystyrene, or glass beads). Any method known in the art can be used to attach the enzyme polypeptide (or polynucleotide) or test compound to a solid support, including use of covalent and non-covalent linkages, passive absorption, or pairs of binding moieties attached respectively to the polypeptide (or polynucleotide) or test compound and the solid support. Test compounds are preferably bound to the solid support in an array, so that the location of individual test compounds can be tracked. Binding of a test compound to a CRIK polypeptide (or polynucleotide) can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and microcentrifuge tubes.

In one embodiment, the CRIK polypeptide is a fusion protein comprising a domain that allows the CRIK polypeptide to be bound to a solid support. For example, glutathione-S-transferase fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, Mo.) or glutathione derivatized microtiter plates, which are then combined with the test compound or the test compound and the non-adsorbed CRIK polypeptide; the mixture is then incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components. Binding of the interactants can be determined either directly or indirectly, as described above. Alternatively, the complexes can be dissociated from the solid support before binding is determined.

Other techniques for immobilizing proteins or polynucleotides on a solid support also can be used in the screening assays of the invention. For example, either a CRIK polypeptide (or polynucleotide) or a test compound can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated CRIK polypeptides (or polynucleotides) or test compounds can be prepared from biotin-NHS(N-hydroxy-succinimide) using techniques well known in the art (*e.g.*, biotinylation kit, Pierce

Chemicals, Rockford, Ill.) and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies which specifically bind to a CRIK polypeptide, polynucleotide, or a test compound, but which do not interfere with a desired binding site, such as the active site of the CRIK polypeptide, can be
5 derivativized to the wells of the plate. Unbound target or protein can be trapped in the wells by antibody conjugation.

Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using
10 antibodies which specifically bind to the CRIK polypeptide or test compound, enzyme-linked assays which rely on detecting an activity of the CRIK polypeptide, and SDS gel electrophoresis under non-reducing conditions.

Screening for test compounds which bind to a CRIK polypeptide or polynucleotide
15 also can be carried out in an intact cell. Any cell which comprises a CRIK polypeptide or polynucleotide can be used in a cell-based assay system. A CRIK polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Binding of the test compound to a CRIK polypeptide or polynucleotide is determined as described above.

20

Enzyme Assays

Test compounds can be tested for the ability to increase or decrease the enzymatic activity of a human CRIK polypeptide. Enzymatic activity can be measured, for
25 example, as described in Di Cunto *et al.*, J Biol Chem. 1998 Nov 6;273(45):29706-11.

Enzyme assays can be carried out after contacting either a purified CRIK polypeptide, a cell membrane preparation, or an intact cell with a test compound. A test
30 compound that decreases an enzymatic activity of a CRIK polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% is identified as

a potential therapeutic agent for decreasing CRIK activity. A test compound which increases an enzymatic activity of a human CRIK polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% is identified as a potential therapeutic agent for increasing human CRIK activity.

5

Gene Expression

In another embodiment, test compounds that increase or decrease CRIK gene expression are identified. A CRIK polynucleotide is contacted with a test compound, and the expression of an RNA or polypeptide product of the CRIK polynucleotide is determined. The level of expression of appropriate mRNA or polypeptide in the presence of the test compound is compared to the level of expression of mRNA or polypeptide in the absence of the test compound. The test compound can then be identified as a modulator of expression based on this comparison. For example, when expression of mRNA or polypeptide is greater in the presence of the test compound than in its absence, the test compound is identified as a stimulator or enhancer of the mRNA or polypeptide expression. Alternatively, when expression of the mRNA or polypeptide is less in the presence of the test compound than in its absence, the test compound is identified as an inhibitor of the mRNA or polypeptide expression.

20

The level of CRIK mRNA or polypeptide expression in the cells can be determined by methods well known in the art for detecting mRNA or polypeptide. Either qualitative or quantitative methods can be used. The presence of polypeptide products of a CRIK polynucleotide can be determined, for example, using a variety of techniques known in the art, including immunochemical methods such as radioimmunoassay, Western blotting, and immunohistochemistry. Alternatively, polypeptide synthesis can be determined *in vivo*, in a cell culture, or in an *in vitro* translation system by detecting incorporation of labeled amino acids into a CRIK polypeptide.

30

Such screening can be carried out either in a cell-free assay system or in an intact cell. Any cell that expresses a CRIK polynucleotide can be used in a cell-based assay system. The CRIK polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Either a primary culture or an established cell line, such as CHO or human embryonic kidney 293 cells, can be used.

Pharmaceutical Compositions

The invention also provides pharmaceutical compositions that can be administered to a patient to achieve a therapeutic effect. Pharmaceutical compositions of the invention can comprise, for example, a CRIK polypeptide, CRIK polynucleotide, ribozymes or antisense oligonucleotides, antibodies which specifically bind to a CRIK polypeptide, or mimetics, activators, or inhibitors of a CRIK polypeptide activity. The compositions can be administered alone or in combination with at least one other agent, such as stabilizing compound, which can be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions can be administered to a patient alone, or in combination with other agents, drugs or hormones.

In addition to the active ingredients, these pharmaceutical compositions can contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries that facilitate processing of the active compounds into preparations which can be used pharmaceutically. Pharmaceutical compositions of the invention can be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, parenteral, topical, sublingual, or rectal means. Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be

- 46 -

formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

5 Pharmaceutical preparations for oral use can be obtained through combination of active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxy-
10 propylmethyl-cellulose, or sodium carboxymethylcellulose; gums including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents can be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

15 Dragee cores can be used in conjunction with suitable coatings, such as concentrated sugar solutions, which also can contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments can be added to the tablets or dragee coatings for product identification or to characterize the quantity
20 of active compound, *i.e.*, dosage.

Pharmaceutical preparations that can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a
25 filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds can be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

30 Pharmaceutical formulations suitable for parenteral administration can be formulated in aqueous solutions, preferably in physiologically compatible buffers such as

- 47 -

Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions can contain substances that increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds can be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Non-lipid polycationic amino polymers also can be used for delivery. Optionally, the suspension also can contain suitable stabilizers or agents that increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention can be manufactured in a manner that is known in the art, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes. The pharmaceutical composition can be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation can be a lyophilized powder which can contain any or all of the following: 1-50 mM histidine, 0.1%-2% sucrose, and 2-7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

Further details on techniques for formulation and administration can be found in the latest edition of REMINGTON'S PHARMACEUTICAL SCIENCES (Maack Publishing Co., Easton, Pa.). After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. Such labeling would include amount, frequency, and method of administration.

Therapeutic Indications and Methods

Human CRIK can be regulated to treat obesity, CNS disorders, and COPD.

5 Obesity. Obesity and overweight are defined as an excess of body fat relative to lean
body mass. An increase in caloric intake or a decrease in energy expenditure or both
can bring about this imbalance leading to surplus energy being stored as fat. Obesity
is associated with important medical morbidities and an increase in mortality. The
causes of obesity are poorly understood and may be due to genetic factors,
10 environmental factors or a combination of the two to cause a positive energy balance.
In contrast, anorexia and cachexia are characterized by an imbalance in energy intake
versus energy expenditure leading to a negative energy balance and weight loss.
Agents that either increase energy expenditure and/or decrease energy intake,
absorption or storage would be useful for treating obesity, overweight, and associated
15 comorbidities. Agents that either increase energy intake and/or decrease energy
expenditure or increase the amount of lean tissue would be useful for treating
cachexia, anorexia and wasting disorders.

This gene, translated proteins and agents which modulate this gene or portions of the
20 gene or its products are useful for treating obesity, overweight, anorexia, cachexia,
wasting disorders, appetite suppression, appetite enhancement, increases or decreases
in satiety, modulation of body weight, and/or other eating disorders such as bulimia.
Also this gene, translated proteins and agents which modulate this gene or portions of
the gene or its products are useful for treating obesity/overweight-associated
25 comorbidities including hypertension, type 2 diabetes, coronary artery disease,
hyperlipidemia, stroke, gallbladder disease, gout, osteoarthritis, sleep apnea and
respiratory problems, some types of cancer including endometrial, breast, prostate,
and colon cancer, thrombotic disease, polycystic ovarian syndrome, reduced fertility,
complications of pregnancy, menstrual irregularities, hirsutism, stress incontinence,
30 and depression.

The hypothalamus has been considered as the feeding control center. Many neuropeptides, hormones, neurotransmitters, etc. that play important roles in the control of energy homeostasis have been identified in the hypothalamus. See *J. Lip. Res.* 40, 1735-46, 1999; *Pharm. Rev.* 52, 35-61, 2000. Leptin signaling pathway, MC4, and 5-HT2C systems in the hypothalamus play critical roles in the control of body weight homeostasis. Therefore, a gene selectively expressed in the hypothalamus, such as the human CRIK of the invention, is a potential obesity target.

CNS disorders. Central and peripheral nervous system disorders also can be treated, such as primary and secondary disorders after brain injury, disorders of mood, anxiety disorders, disorders of thought and volition, disorders of sleep and wakefulness, diseases of the motor unit, such as neurogenic and myopathic disorders, neurodegenerative disorders such as Alzheimer's and Parkinson's disease, and processes of peripheral and chronic pain. Pain that is associated with CNS disorders also can be treated by regulating the activity of human CRIK. Pain which can be treated includes that associated with central nervous system disorders, such as multiple sclerosis, spinal cord injury, sciatica, failed back surgery syndrome, traumatic brain injury, epilepsy, Parkinson's disease, post-stroke, and vascular lesions in the brain and spinal cord (e.g., infarct, hemorrhage, vascular malformation). Non-central neuropathic pain includes that associated with post mastectomy pain, reflex sympathetic dystrophy (RSD), trigeminal neuralgia-radioculopathy, post-surgical pain, HIV/AIDS related pain, cancer pain, metabolic neuropathies (e.g., diabetic neuropathy, vasculitic neuropathy secondary to connective tissue disease), paraneoplastic polyneuropathy associated, for example, with carcinoma of lung, or leukemia, or lymphoma, or carcinoma of prostate, colon or stomach, trigeminal neuralgia, cranial neuralgias, and post-herpetic neuralgia. Pain associated with cancer and cancer treatment also can be treated, as can headache pain (for example, migraine with aura, migraine without aura, and other migraine disorders), episodic and chronic tension-type headache, tension-type like headache, cluster headache, and chronic paroxysmal hemicrania.

COPD. Chronic obstructive pulmonary (or airways) disease (COPD) is a condition defined physiologically as airflow obstruction that generally results from a mixture of emphysema and peripheral airway obstruction due to chronic bronchitis (Senior & Shapiro, *Pulmonary Diseases and Disorders*, 3d ed., New York, McGraw-Hill, 1998, pp. 659-681, 1998; Barnes, *Chest* 117, 10S-14S, 2000). Emphysema is characterized by destruction of alveolar walls leading to abnormal enlargement of the air spaces of the lung. Chronic bronchitis is defined clinically as the presence of chronic productive cough for three months in each of two successive years. In COPD, airflow obstruction is usually progressive and is only partially reversible. By far the most important risk factor for development of COPD is cigarette smoking, although the disease does occur in non-smokers.

Chronic inflammation of the airways is a key pathological feature of COPD (Senior & Shapiro, 1998). The inflammatory cell population comprises increased numbers of macrophages, neutrophils, and CD8⁺ lymphocytes. Inhaled irritants, such as cigarette smoke, activate macrophages which are resident in the respiratory tract, as well as epithelial cells leading to release of chemokines (*e.g.*, interleukin-8) and other chemotactic factors. These chemotactic factors act to increase the neutrophil/monocyte trafficking from the blood into the lung tissue and airways. Neutrophils and monocytes recruited into the airways can release a variety of potentially damaging mediators such as proteolytic enzymes and reactive oxygen species. Matrix degradation and emphysema, along with airway wall thickening, surfactant dysfunction, and mucus hypersecretion, all are potential sequelae of this inflammatory response that lead to impaired airflow and gas exchange.

Protein kinases are signal transducing enzymes that phosphorylate proteins, including other kinases, and, along with protein phosphatases, regulate the level and extent of protein phosphorylation and activation. Intracellular signalling pathways have important roles in inflammatory processes. These pathways may be activated by cytokines, oxidant stress and other inflammatory mediators (reviewed in Kyraikis and Avruch, 1996 and 2001). For example, the pro-inflammatory cytokines, tumor

necrosis factor α (TNF α) and interleukin-1 activate the protein ser/thr kinases c-Jun-NH2-terminal kinase (JNK) and p38 mitogen-activated protein (MAP) kinase, leading to activation of AP-1 and IKB kinase (IKK), which, in turn, leads to activation of the transcription factor NFKB. Activation of NFKB is required for the transcription of several pro-inflammatory molecules, including interleukin-8 and ICAM-1. Enzymes of the MAP kinase class may also act to increase cytokine production by stabilization of mRNA (Winzen et al., 1999).

Inhibition of specific protein kinases has been shown to elicit anti-inflammatory effects. For example, the accumulation of polymorphonuclear leukocytes in murine lung following intratracheal administration of bacterial lipopolysaccharide can be blocked by inhibition of p38 MAP kinase (Nick, et al. 2000). As a further example, aerosol delivery to rat lungs of antisense oligodeoxynucleotides to syk kinase mRNA, suppressed nitric oxide and TNF α production from alveolar macrophages stimulated with IgG-anti-IgG complexes (Stenton et al. 2000). Protein kinase subtypes are therefore attractive therapeutic targets for the attenuation of the inflammatory response in COPD. See Kyriakis, J.M. and Avruch J. Sounding the alarm: protein kinase cascades activated by stress and inflammation. *J Biol Chem* 1996, 271:24313-6; Kyriakis, J.M. and Avruch, J. Mammalian mitogen-activated protein kinase signal transduction pathways activated by stress and inflammation. *J. Physiol. Rev.* 2001, 81:807-69; Winzen, R., Kracht, M., Ritter, B., Wilhelm, A., Chen C.A., Shyu, A., Müller, M., Gaestel, M., Resch, K., and Holtmann, H. The p38 MAP kinase pathway signals for cytokine-induced mRNA stabilization via MAP kinase-activated protein kinase 2 and an AU-rich region-targeted mechanism. *EMBO J.* 1999, 18: 4969-4980; Nick, J.A., Young, S.K., Brown, K.K., Avdi, N.J., Arndt, P.G., Suratt, B.T., Janes, M.S., Henson, P.M., Worthen, G.S. Role of p38 mitogen-activated protein kinase in a murine model of pulmonary inflammation. *J Immunol.* 2000, 164:2151-9; and Stenton, G.R., Kim, M.K., Nohara, O., Chen, C.F., Hirji, N., Wills, F.L., Gilchrist, M., Hwang, P.H., Park, J.G., Finlay, W., Jones, R.L., Befus, A.D., Schreiber, A.D. Aerosolized Syk antisense suppresses Syk expression,

mediator release from macrophages, and pulmonary inflammation. J Immunol 2000, 164:3790-7.

5 This invention further pertains to the use of novel agents identified by the screening assays described above. Accordingly, it is within the scope of this invention to use a test compound identified as described herein in an appropriate animal model. For example, an agent identified as described herein (e.g., a modulating agent, an antisense nucleic acid molecule, a specific antibody, ribozyme, or a CRIK polypeptide binding molecule) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

15 A reagent which affects CRIK activity can be administered to a human cell, either *in vitro* or *in vivo*, to reduce CRIK activity. The reagent preferably binds to an expression product of a human CRIK gene. If the expression product is a protein, the reagent is preferably an antibody. For treatment of human cells *ex vivo*, an antibody can be added to a preparation of stem cells that have been removed from the body. The cells can then be replaced in the same or another human body, with or without clonal propagation, as is known in the art.

25 In one embodiment, the reagent is delivered using a liposome. Preferably, the liposome is stable in the animal into which it has been administered for at least about 30 minutes, more preferably for at least about 1 hour, and even more preferably for at least about 24 hours. A liposome comprises a lipid composition that is capable of targeting a reagent, particularly a polynucleotide, to a particular site in an animal, such as a human. Preferably, the lipid composition of the liposome is capable of targeting to a specific organ of an animal, such as the lung, liver, spleen, heart brain, lymph nodes, and skin.

A liposome useful in the present invention comprises a lipid composition that is capable of fusing with the plasma membrane of the targeted cell to deliver its contents to the cell. Preferably, the transfection efficiency of a liposome is about 5 0.5 μg of DNA per 16 nmole of liposome delivered to about 10^6 cells, more preferably about 1.0 μg of DNA per 16 nmole of liposome delivered to about 10^6 cells, and even more preferably about 2.0 μg of DNA per 16 nmol of liposome delivered to about 10^6 cells. Preferably, a liposome is between about 100 and 500 nm, more preferably between about 150 and 450 nm, and even more preferably 10 between about 200 and 400 nm in diameter.

Suitable liposomes for use in the present invention include those liposomes standardly used in, for example, gene delivery methods known to those of skill in the art. More preferred liposomes include liposomes having a polycationic lipid 15 composition and/or liposomes having a cholesterol backbone conjugated to polyethylene glycol. Optionally, a liposome comprises a compound capable of targeting the liposome to a particular cell type, such as a cell-specific ligand exposed on the outer surface of the liposome.

20 Complexing a liposome with a reagent such as an antisense oligonucleotide or ribozyme can be achieved using methods that are standard in the art (see, for example, U.S. Patent 5,705,151). Preferably, from about 0.1 μg to about 10 μg of polynucleotide is combined with about 8 nmol of liposomes, more preferably from about 0.5 μg to about 5 μg of polynucleotides are combined with about 8 nmol lipo- 25 somes, and even more preferably about 1.0 μg of polynucleotides is combined with about 8 nmol liposomes.

In another embodiment, antibodies can be delivered to specific tissues *in vivo* using receptor-mediated targeted delivery. Receptor-mediated DNA delivery techniques 30 are taught in, for example, Findeis *et al. Trends in Biotechnol.* 11, 202-05 (1993); Chiou *et al.*, GENE THERAPEUTICS: METHODS AND APPLICATIONS OF DIRECT GENE

TRANSFER (J.A. Wolff, ed.) (1994); Wu & Wu, *J. Biol. Chem.* 263, 621-24 (1988); Wu *et al.*, *J. Biol. Chem.* 269, 542-46 (1994); Zenke *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 87, 3655-59 (1990); Wu *et al.*, *J. Biol. Chem.* 266, 338-42 (1991).

5 Determination of a Therapeutically Effective Dose

10 The determination of a therapeutically effective dose is well within the capability of those skilled in the art. A therapeutically effective dose refers to that amount of active ingredient which increases or decreases CRIK activity relative to the CRIK activity which occurs in the absence of the therapeutically effective dose.

15 For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays or in animal models, usually mice, rabbits, dogs, or pigs. The animal model also can be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

20 Therapeutic efficacy and toxicity, *e.g.*, ED₅₀ (the dose therapeutically effective in 50% of the population) and LD₅₀ (the dose lethal to 50% of the population), can be determined by standard pharmaceutical procedures in cell cultures or experimental animals. The dose ratio of toxic to therapeutic effects is the therapeutic index, and it can be expressed as the ratio, LD₅₀/ED₅₀.

25 Pharmaceutical compositions that exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED₅₀ with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active ingredient or to maintain the desired effect. Factors that can be taken into account include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions can be administered every 3 to 4 days, every week, or once every two weeks depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts can vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

If the reagent is a single-chain antibody, polynucleotides encoding the antibody can be constructed and introduced into a cell either *ex vivo* or *in vivo* using well-established techniques including, but not limited to, transferrin-polycation-mediated DNA transfer, transfection with naked or encapsulated nucleic acids, liposome-mediated cellular fusion, intracellular transportation of DNA-coated latex beads, protoplast fusion, viral infection, electroporation, "gene gun," and DEAE- or calcium phosphate-mediated transfection.

Effective *in vivo* dosages of an antibody are in the range of about 5 µg to about 50 µg/kg, about 50 µg to about 5 mg/kg, about 100 µg to about 500 µg/kg of patient body weight, and about 200 to about 250 µg/kg of patient body weight. For administration of polynucleotides encoding single-chain antibodies, effective *in vivo* dosages are in the range of about 100 ng to about 200 ng, 500 ng to about 50 mg,

about 1 μ g to about 2 mg, about 5 μ g to about 500 μ g, and about 20 μ g to about 100 μ g of DNA.

5 If the expression product is mRNA, the reagent is preferably an antisense oligonucleotide or a ribozyme. Polynucleotides that express antisense oligonucleotides or ribozymes can be introduced into cells by a variety of methods, as described above.

10 Preferably, a reagent reduces expression of a CRIK gene or the activity of a CRIK polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the reagent. The effectiveness of the mechanism chosen to decrease the level of expression of a CRIK gene or the activity of a CRIK polypeptide can be assessed using methods well known in the art, such as hybridization of nucleotide probes to CRIK-specific mRNA, quantitative RT-PCR, immunologic detection of a CRIK polypeptide, or measurement of CRIK activity.

15 In any of the embodiments described above, any of the pharmaceutical compositions of the invention can be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy can be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents can act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

25 Any of the therapeutic methods described above can be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

Diagnostic Methods

Human CRIK also can be used in diagnostic assays for detecting diseases and abnormalities or susceptibility to diseases and abnormalities related to the presence of mutations in the nucleic acid sequences that encode the enzyme. For example, differences can be determined between the cDNA or genomic sequence encoding CRIK in individuals afflicted with a disease and in normal individuals. If a mutation is observed in some or all of the afflicted individuals but not in normal individuals, then the mutation is likely to be the causative agent of the disease.

10

Sequence differences between a reference gene and a gene having mutations can be revealed by the direct DNA sequencing method. In addition, cloned DNA segments can be employed as probes to detect specific DNA segments. The sensitivity of this method is greatly enhanced when combined with PCR. For example, a sequencing primer can be used with a double-stranded PCR product or a single-stranded template molecule generated by a modified PCR. The sequence determination is performed by conventional procedures using radiolabeled nucleotides or by automatic sequencing procedures using fluorescent tags.

15

Genetic testing based on DNA sequence differences can be carried out by detection of alteration in electrophoretic mobility of DNA fragments in gels with or without denaturing agents. Small sequence deletions and insertions can be visualized, for example, by high resolution gel electrophoresis. DNA fragments of different sequences can be distinguished on denaturing formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions according to their specific melting or partial melting temperatures (*see, e.g., Myers et al., Science 230, 1242, 1985*). Sequence changes at specific locations can also be revealed by nuclease protection assays, such as RNase and S 1 protection or the chemical cleavage method (*e.g., Cotton et al., Proc. Natl. Acad. Sci. USA 85, 4397-4401, 1985*). Thus, the detection of a specific DNA sequence can be performed by methods such as hybridization, RNase protection, chemical cleavage, direct DNA

20

25

30

sequencing or the use of restriction enzymes and Southern blotting of genomic DNA. In addition to direct methods such as gel-electrophoresis and DNA sequencing, mutations can also be detected by *in situ* analysis.

5 Altered levels of CRIK also can be detected in various tissues. Assays used to detect levels of the receptor polypeptides in a body sample, such as blood or a tissue biopsy, derived from a host are well known to those of skill in the art and include radioimmunoassays, competitive binding assays, Western blot analysis, and ELISA assays.

10

All patents and patent applications cited in this disclosure are expressly incorporated herein by reference. The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following specific examples, which are provided for purposes of illustration only and are not
15 intended to limit the scope of the invention.

EXAMPLE 1

Detection of human citron rho/rac-interacting kinase activity

20 Subconfluent COS7 cells in 10-cm dishes are transiently transfected by the DEAE-dextran/chloroquine method with 10 µg of FLAG-SEQ ID NO: 1 vector. Cells are harvested 48 h after transfection. Immunoblotting is performed, and cells are probed with anti-FLAG M2 antibodies (Eastman Kodak Co.) Blots are developed using horseradish peroxidase-conjugated secondary antibodies and ECL detection system
25 (Amersham Pharmacia Biotech). In vitro kinase assays are performed by incubating immune complexes in 50 µl of kinase buffer (50 mM HEPES, pH 7.4, 5 mM MgCl₂, 3 mM MnCl₂, 1mM dithiothreitol), in the presence or absence of 5 µg of histone H1 or myelin basic protein, plus 0.1 mM ATP and 10 mCi of [gamma-32P] ATP (6000 Ci/mM, NEN Life Science Products) for 30 min at 30 °C. The products are
30 analyzed by 5% or 12.5% SDS-PAGE followed by autoradiography. For immunoprecipitation of metabolically labeled proteins, primary keratinocytes are

incubated with 0.1 mCi/ml [³⁵S]methionine (Expre35S; NEN Life Science Products) for 4 h in methionine-free medium in the presence of serum. Immunoprecipitated proteins are separated on a 5% SDS-PAGE gel and visualized by autoradiography. It is shown that the polypeptide of SEQ ID NO: 2 has a human citron rho/rac-interacting kinase-short kinase activity.

EXAMPLE 2

Expression of recombinant human CRIK

The *Pichia pastoris* expression vector pPICZB (Invitrogen, San Diego, CA) is used to produce large quantities of recombinant human CRIK polypeptides in yeast. The CRIK-encoding DNA sequence is derived from SEQ ID NO:1. Before insertion into vector pPICZB, the DNA sequence is modified by well known methods in such a way that it contains at its 5'-end an initiation codon and at its 3'-end an enterokinase cleavage site, a His6 reporter tag and a termination codon. Moreover, at both termini recognition sequences for restriction endonucleases are added and after digestion of the multiple cloning site of pPICZ B with the corresponding restriction enzymes the modified DNA sequence is ligated into pPICZB. This expression vector is designed for inducible expression in *Pichia pastoris*, driven by a yeast promoter. The resulting pPICZ/md-His6 vector is used to transform the yeast.

The yeast is cultivated under usual conditions in 5 liter shake flasks and the recombinantly produced protein isolated from the culture by affinity chromatography (Ni-NTA-Resin) in the presence of 8 M urea. The bound polypeptide is eluted with buffer, pH 3.5, and neutralized. Separation of the polypeptide from the His6 reporter tag is accomplished by site-specific proteolysis using enterokinase (Invitrogen, San Diego, CA) according to manufacturer's instructions. Purified human CRIK polypeptide is obtained.

EXAMPLE 3*Identification of test compounds that bind to CRIK polypeptides*

Purified CRIK polypeptides comprising a glutathione-S-transferase protein and
5 absorbed onto glutathione-derivatized wells of 96-well microtiter plates are contacted
with test compounds from a small molecule library at pH 7.0 in a physiological
buffer solution. Human CRIK polypeptides comprise the amino acid sequence
shown in SEQ ID NO:2. The test compounds comprise a fluorescent tag. The
10 samples are incubated for 5 minutes to one hour. Control samples are incubated in
the absence of a test compound.

The buffer solution containing the test compounds is washed from the wells.
Binding of a test compound to a CRIK polypeptide is detected by fluorescence
measurements of the contents of the wells. A test compound that increases the
15 fluorescence in a well by at least 15% relative to fluorescence of a well in which a
test compound is not incubated is identified as a compound which binds to a CRIK
polypeptide.

EXAMPLE 4*Identification of a test compound which decreases CRIK gene expression*

A test compound is administered to a culture of human cells transfected with a CRIK
expression construct and incubated at 37 °C for 10 to 45 minutes. A culture of the
same type of cells that have not been transfected is incubated for the same time
25 without the test compound to provide a negative control.

RNA is isolated from the two cultures as described in Chirgwin *et al.*, *Biochem.* 18,
5294-99, 1979). Northern blots are prepared using 20 to 30 µg total RNA and
hybridized with a ³²P-labeled CRIK-specific probe at 65 °C in Express-hyb
30 (CLONTECH). The probe comprises at least 11 contiguous nucleotides selected
from the complement of SEQ ID NO:1. A test compound that decreases the CRIK-

specific signal relative to the signal obtained in the absence of the test compound is identified as an inhibitor of CRIK gene expression.

EXAMPLE 5

5 *Identification of a test compound which decreases CRIK activity*

A test compound is administered to a culture of human cells transfected with a CRIK expression construct and incubated at 37 °C for 10 to 45 minutes. A culture of the same type of cells that have not been transfected is incubated for the same time
10 without the test compound to provide a negative control. CRIK activity is measured using the method of Di Cunto *et al.*, J Biol Chem. 1998 Nov 6;273(45):29706-11.

A test compound which decreases the CRIK activity of the CRIK relative to the CRIK activity in the absence of the test compound is identified as an inhibitor of
15 CRIK activity.

EXAMPLE 6

Tissue-specific expression of CRIK

20 The qualitative expression pattern of CRIK in various tissues is determined by Reverse Transcription-Polymerase Chain Reaction (RT-PCR).

To demonstrate that CRIK is involved in the disease process of COPD, the initial expression panel consists of RNA samples from respiratory tissues and inflammatory cells relevant to COPD: lung (adult and fetal), trachea, freshly isolated alveolar type
25 II cells, cultured human bronchial epithelial cells, cultured small airway epithelial cells, cultured bronchial smooth muscle cells, cultured H441 cells (Clara-like), freshly isolated neutrophils and monocytes, and cultured monocytes (macrophage-like). Body map profiling also is carried out, using total RNA panels purchased from
30 Clontech. The tissues are adrenal gland, bone marrow, brain, colon, heart, kidney,

- 62 -

liver, lung, mammary gland, pancreas, prostate, salivary gland, skeletal muscle, small intestine, spleen, stomach, testis, thymus, trachea, thyroid, and uterus.

To demonstrate that CRIK is involved in CNS disorders, the following tissues are
5 screened: fetal and adult brain, muscle, heart, lung, kidney, liver, thymus, testis, colon, placenta, trachea, pancreas, kidney, gastric mucosa, colon, liver, cerebellum, skin, cortex (Alzheimer's and normal), hypothalamus, cortex, amygdala, cerebellum, hippocampus, choroid, plexus, thalamus, and spinal cord.

10 To demonstrate that CRIK is involved in the disease process of obesity, expression is determined in the following tissues: subcutaneous adipose tissue, mesenteric adipose tissue, adrenal gland, bone marrow, brain (cerebellum, spinal cord, cerebral cortex, caudate, medulla, substantia nigra, and putamen), colon, fetal brain, heart, kidney, liver, lung, mammary gland, pancreas, placenta, prostate, salivary gland, skeletal
15 muscle small intestine, spleen, stomach, testes, thymus, thyroid trachea, and uterus. Neuroblastoma cell lines SK-Nr-Be (2), Hr, Sk-N-As, HTB-10, IMR-32, SNSY-5Y, T3, SK-N-D2, D283, DAOY, CHP-2, U87MG, BE(2)C, T986, KANTS, MO59K, CHP234, C6 (rat), SK-N-F1, SK-PU-DW, PFSK-1, BE(2)M17, and MCIXC also are tested for CRIK expression. As a final step, the expression of CRIK in cells derived
20 from normal individuals with the expression of cells derived from obese individuals is compared.

Quantitative expression profiling. Quantitative expression profiling is performed by the form of quantitative PCR analysis called "kinetic analysis" firstly described in
25 Higuchi *et al.*, *BioTechnology* 10, 413-17, 1992, and Higuchi *et al.*, *BioTechnology* 11, 1026-30, 1993. The principle is that at any given cycle within the exponential phase of PCR, the amount of product is proportional to the initial number of template copies.

30 If the amplification is performed in the presence of an internally quenched fluorescent oligonucleotide (TaqMan probe) complementary to the target sequence,

- 63 -

the probe is cleaved by the 5'-3' endonuclease activity of Taq DNA polymerase and a fluorescent dye released in the medium (Holland *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 88, 7276-80, 1991). Because the fluorescence emission will increase in direct proportion to the amount of the specific amplified product, the exponential growth phase of PCR product can be detected and used to determine the initial template concentration (Heid *et al.*, *Genome Res.* 6, 986-94, 1996, and Gibson *et al.*, *Genome Res.* 6, 995-1001, 1996).

The amplification of an endogenous control can be performed to standardize the amount of sample RNA added to a reaction. In this kind of experiment, the control of choice is the 18S ribosomal RNA. Because reporter dyes with differing emission spectra are available, the target and the endogenous control can be independently quantified in the same tube if probes labeled with different dyes are used.

All "real time PCR" measurements of fluorescence are made in the ABI Prism 7700.

RNA extraction and cDNA preparation. Total RNA from the tissues listed above are used for expression quantification. RNAs labeled "from autopsy" were extracted from autoptic tissues with the TRIzol reagent (Life Technologies, MD) according to the manufacturer's protocol.

Fifty µg of each RNA were treated with DNase I for 1 hour at 37°C in the following reaction mix: 0.2 U/µl RNase-free DNase I (Roche Diagnostics, Germany); 0.4 U/µl RNase inhibitor (PE Applied Biosystems, CA); 10 mM Tris-HCl pH 7.9; 10mM MgCl₂; 50 mM NaCl; and 1 mM DTT.

After incubation, RNA is extracted once with 1 volume of phenol:chloroform:isoamyl alcohol (24:24:1) and once with chloroform, and precipitated with 1/10 volume of 3 M NaAcetate, pH5.2, and 2 volumes of ethanol.

Fifty μg of each RNA from the autaptic tissues are DNase treated with the DNA-free kit purchased from Ambion (Ambion, TX). After resuspension and spectrophotometric quantification, each sample is reverse transcribed with the TaqMan Reverse Transcription Reagents (PE Applied Biosystems, CA) according to the manufacturer's protocol. The final concentration of RNA in the reaction mix is 200ng/ μL . Reverse transcription is carried out with 2.5 μM of random hexamer primers.

TaqMan quantitative analysis. Specific primers and probe are designed according to the recommendations of PE Applied Biosystems; the probe can be labeled at the 5' end FAM (6-carboxy-fluorescein) and at the 3' end with TAMRA (6-carboxy-tetramethyl-rhodamine). Quantification experiments are performed on 10 ng of reverse transcribed RNA from each sample. Each determination is done in triplicate.

Total cDNA content is normalized with the simultaneous quantification (multiplex PCR) of the 18S ribosomal RNA using the Pre-Developed TaqMan Assay Reagents (PDAR) Control Kit (PE Applied Biosystems, CA).

The assay reaction mix is as follows: 1X final TaqMan Universal PCR Master Mix (from 2X stock) (PE Applied Biosystems, CA); 1X PDAR control – 18S RNA (from 20X stock); 300 nM forward primer; 900 nM reverse primer; 200 nM probe; 10 ng cDNA; and water to 25 μL .

Each of the following steps are carried out once: pre PCR, 2 minutes at 50 °C, and 10 minutes at 95 °C. The following steps are carried out 40 times: denaturation, 15 seconds at 95 °C, annealing/extension, 1 minute at 60 °C.

The experiment is performed on an ABI Prism 7700 Sequence Detector (PE Applied Biosystems, CA). At the end of the run, fluorescence data acquired during PCR are processed as described in the ABI Prism 7700 user's manual in order to achieve better background subtraction as well as signal linearity with the starting target quantity.

EXAMPLE 7*Identification of test compound efficacy in a COPD animal model*

5 Guinea pigs are exposed on a single occasion to tobacco smoke for 50 minutes. Animals are sacrificed between 10 minutes and 24 hour following the end of the exposure and their lungs placed in RNeasyTM. The lung tissue is homogenized, and total RNA was extracted using a Qiagen RNeasyTM Maxi kit. Molecular Probes RiboGreenTM RNA quantitation method is used to quantify the amount of RNA in
10 each sample.

Total RNA is reverse transcribed, and the resultant cDNA is used in a real-time polymerase chain reaction (PCR). The cDNA is added to a solution containing the sense and anti-sense primers and the 6-carboxy-tetramethyl-rhodamine labelled probe
15 of the CRIK gene. Cyclophilin is used as the housekeeping gene. The expression of the CRIK gene is measured using the TaqMan real-time PCR system that generates an amplification curve for each sample. From this curve a threshold cycle value is calculated: the fractional cycle number at which the amount of amplified target reaches a fixed threshold. A sample containing many copies of the CRIK gene will
20 reach this threshold earlier than a sample containing fewer copies. The threshold is set at 0.2, and the threshold cycle C_T is calculated from the amplification curve. The C_T value for the CRIK gene is normalized using the C_T value for the housekeeping gene.

25 Expression of the CRIK gene is increased by at least 3-fold between 10 minutes and 3 hours post tobacco smoke exposure compared to air exposed control animals.

Test compounds are evaluated as follows. Animals are pre-treated with a test compound between 5 minutes and 1 hour prior to the tobacco smoke exposure and
30 they are then sacrificed up to 3 hours after the tobacco smoke exposure has been completed. Control animals are pre-treated with the vehicle of the test compound via

- 66 -

the route of administration chosen for the test compound. A test compound that reduces the tobacco smoke induced upregulation of CRIK gene relative to the expression seen in vehicle treated tobacco smoke exposed animals is identified as an inhibitor of CRIK gene expression.

5

EXAMPLE 8

Expression of human citron rho/rac-interacting kinase

10 Total RNA used for Taqman quantitative analysis were either purchased (Clontech, CA) or extracted from tissues using TRIzol reagent (Life Technologies, MD) according to a modified vendor protocol which utilizes the Rneasy protocol (Qiagen, CA). One hundred μ g of each RNA were treated with DNase I using RNase free- DNase (Qiagen, CA) for use with RNeasy or QiaAmp columns.

15 After elution and quantitation with Ribogreen (Molecular Probes Inc., OR), each sample was reverse transcribed using the GibcoBRL Superscript II First Strand Synthesis System for RT-PCR according to vendor protocol (Life Technologies, MD). The final concentration of RNA in the reaction mix was 50ng/ μ L. Reverse transcription was performed with 50 ng of random hexamers.

20

Specific primers and probe were designed according to PE Applied Biosystems' Primer Express program recommendations and are listed below:

forward primer: 5'-(TCCAATTTTGATGAACCAGAGAAG)-3'

25 reverse primer: 5'-(AACCCACAAACGGCAGTT)-3'

probe: SYBR Green

Quantitation experiments were performed on 25 ng of reverse transcribed RNA from each sample. 18S ribosomal RNA was measured as a control using the Pre-
30 Developed TaqMan Assay Reagents (PDAR)(PE Applied Biosystems, CA). The assay reaction mix was as follows:

- 67 -

- | | | |
|----|---------------------------------------|-------|
| | | final |
| | TaqMan SYBR Green PCR Master Mix (2x) | 1x |
| | (PE Applied Biosystems, CA) | |
| 5 | Forward primer | 300nM |
| | Reverse primer | 300nM |
| | cDNA | 25ng |
| | Water to 25uL | |
| | PCR conditions: | |
| 10 | Once: 2' minutes at 50° C | |
| | 10 minutes at 95°C | |
| | 40cycles: 15 sec.at 95°C | |
| | 1 minute at 60°C | |
- 15 The experiment was performed on an ABI Prism 7700 Sequence Detector (PE Applied Biosystems, CA). At the end of the run, fluorescence data acquired during PCR were processed as described in the ABI Prism 7700 user's manual. Fold change was calculated using the delta-delta CT method with normalization to the 18S values. Relative expression was calculated by normalizing to 18s (D Ct), then making the
- 20 highest expressing tissue 100 and everything else relative to it. Copy number conversion was performed without normalization using the formula $C_n = 10^{(C_t - 40.007)/-3.623}$.
- The results are shown in FIG. 21.
- 25 Human citron rho/rac-interacting kinase expressed in adipose and skeletal muscle could be regulated to increase insulin sensitivity.

EXAMPLE 9*In vivo testing of compounds/target validation***1. Pain:*****Acute Pain***

5

Acute pain is measured on a hot plate mainly in rats. Two variants of hot plate testing are used: In the classical variant animals are put on a hot surface (52 to 56 °C) and the latency time is measured until the animals show nocifensive behavior, such as stepping or foot licking. The other variant is an increasing temperature hot plate where the experimental animals are put on a surface of neutral temperature. Subsequently this surface is slowly but constantly heated until the animals begin to lick a hind paw. The temperature which is reached when hind paw licking begins is a measure for pain threshold.

15 Compounds are tested against a vehicle treated control group. Substance application is performed at different time points via different application routes (i.v., i.p., p.o., i.t., i.c.v., s.c., intradermal, transdermal) prior to pain testing.

Persistent Pain

20

Persistent pain is measured with the formalin or capsaicin test, mainly in rats. A solution of 1 to 5% formalin or 10 to 100 µg capsaicin is injected into one hind paw of the experimental animal. After formalin or capsaicin application the animals show nocifensive reactions like flinching, licking and biting of the affected paw. The number of nocifensive reactions within a time frame of up to 90 minutes is a measure for intensity of pain.

25 Compounds are tested against a vehicle treated control group. Substance application is performed at different time points via different application routes (i.v., i.p., p.o., i.t., i.c.v., s.c., intradermal, transdermal) prior to formalin or capsaicin administration.

30

Neuropathic Pain

Neuropathic pain is induced by different variants of unilateral sciatic nerve injury mainly in rats. The operation is performed under anesthesia. The first variant of sciatic nerve injury is produced by placing loosely constrictive ligatures around the common sciatic nerve. The second variant is the tight ligation of about the half of the diameter of the common sciatic nerve. In the next variant, a group of models is used in which tight ligations or transections are made of either the L5 and L6 spinal nerves, or the L₅ spinal nerve only. The fourth variant involves an axotomy of two of the three terminal branches of the sciatic nerve (tibial and common peroneal nerves) leaving the remaining sural nerve intact whereas the last variant comprises the axotomy of only the tibial branch leaving the sural and common nerves uninjured. Control animals are treated with a sham operation.

Postoperatively, the nerve injured animals develop a chronic mechanical allodynia, cold allodynia, as well as a thermal hyperalgesia. Mechanical allodynia is measured by means of a pressure transducer (electronic von Frey Anesthesiometer, IITC Inc.-Life Science Instruments, Woodland Hills, SA, USA; Electronic von Frey System, Somedic Sales AB, Hörby, Sweden). Thermal hyperalgesia is measured by means of a radiant heat source (Plantar Test, Ugo Basile, Comerio, Italy), or by means of a cold plate of 5 to 10 °C where the nocifensive reactions of the affected hind paw are counted as a measure of pain intensity. A further test for cold induced pain is the counting of nocifensive reactions, or duration of nocifensive responses after plantar administration of acetone to the affected hind limb. Chronic pain in general is assessed by registering the circadian rhythms in activity (Surjo and Arndt, Universität zu Köln, Cologne, Germany), and by scoring differences in gait (foot print patterns; FOOTPRINTS program, Klapdor et al., 1997. A low cost method to analyze footprint patterns. J. Neurosci. Methods 75, 49-54).

Compounds are tested against sham operated and vehicle treated control groups. Substance application is performed at different time points via different application routes (i.v., i.p., p.o., i.t., i.c.v., s.c., intradermal, transdermal) prior to pain testing.

5 *Inflammatory Pain*

Inflammatory pain is induced mainly in rats by injection of 0.75 mg carrageenan or complete Freund's adjuvant into one hind paw. The animals develop an edema with mechanical allodynia as well as thermal hyperalgesia. Mechanical allodynia is measured by means of a pressure transducer (electronic von Frey Anesthesiometer, IITC Inc.-Life Science Instruments, Woodland Hills, SA, USA). Thermal hyperalgesia is measured by means of a radiant heat source (Plantar Test, Ugo Basile, Comerio, Italy, Paw thermal stimulator, G. Ozaki, University of California, USA). For edema measurement two methods are being used. In the first method, the animals are sacrificed and the affected hindpaws sectioned and weighed. The second method comprises differences in paw volume by measuring water displacement in a plethysmometer (Ugo Basile, Comerio, Italy).

Compounds are tested against uninflamed as well as vehicle treated control groups. Substance application is performed at different time points via different application routes (i.v., i.p., p.o., i.t., i.c.v., s.c., intradermal, transdermal) prior to pain testing.

Diabetic Neuropathic Pain

Rats treated with a single intraperitoneal injection of 50 to 80 mg/kg streptozotocin develop a profound hyperglycemia and mechanical allodynia within 1 to 3 weeks. Mechanical allodynia is measured by means of a pressure transducer (electronic von Frey Anesthesiometer, IITC Inc.-Life Science Instruments, Woodland Hills, SA, USA).

Compounds are tested against diabetic and non-diabetic vehicle treated control groups. Substance application is performed at different time points via different application routes (i.v., i.p., p.o., i.t., i.c.v., s.c., intradermal, transdermal) prior to pain testing.

5

2. Parkinson's disease

6-Hydroxydopamine (6-OH-DA) Lesion

Degeneration of the dopaminergic nigrostriatal and striatopallidal pathways is the central pathological event in Parkinson's disease. This disorder has been mimicked experimentally in rats using single/sequential unilateral stereotaxic injections of 6-OH-DA into the medium forebrain bundle (MFB).

Male Wistar rats (Harlan Winkelmann, Germany), weighing 200±250 g at the beginning of the experiment, are used. The rats are maintained in a temperature- and humidity-controlled environment under a 12 h light/dark cycle with free access to food and water when not in experimental sessions. The following in vivo protocols are approved by the governmental authorities. All efforts are made to minimize animal suffering, to reduce the number of animals used, and to utilize alternatives to in vivo techniques.

Animals are administered pargyline on the day of surgery (Sigma, St. Louis, MO, USA; 50 mg/kg i.p.) in order to inhibit metabolism of 6-OHDA by monoamine oxidase and desmethylimipramine HCl (Sigma; 25 mg/kg i.p.) in order to prevent uptake of 6-OHDA by noradrenergic terminals. Thirty minutes later the rats are anesthetized with sodium pentobarbital (50 mg/kg) and placed in a stereotaxic frame. In order to lesion the DA nigrostriatal pathway 4 µl of 0.01% ascorbic acid-saline containing 8 µg of 6-OHDA HBr (Sigma) are injected into the left medial fore-brain bundle at a rate of 1 µl/min (2.4 mm anterior, 1.49 mm lateral, -2.7 mm ventral to Bregma and the skull surface). The needle is left in place an additional 5 min to allow diffusion to occur.

Stepping Test

Forelimb akinesia is assessed three weeks following lesion placement using a modified stepping test protocol. In brief, the animals are held by the experimenter with one hand fixing the hindlimbs and slightly raising the hind part above the surface. One paw is touching the table, and is then moved slowly sideways (5 s for 1 m), first in the forehand and then in the backhand direction. The number of adjusting steps is counted for both paws in the backhand and forehand direction of movement. The sequence of testing is right paw forehand and backhand adjusting stepping, followed by left paw forehand and backhand directions. The test is repeated three times on three consecutive days, after an initial training period of three days prior to the first testing. Forehand adjusted stepping reveals no consistent differences between lesioned and healthy control animals. Analysis is therefore restricted to backhand adjusted stepping.

Balance Test

Balance adjustments following postural challenge are also measured during the stepping test sessions. The rats are held in the same position as described in the stepping test and, instead of being moved sideways, tilted by the experimenter towards the side of the paw touching the table. This maneuver results in loss of balance and the ability of the rats to regain balance by forelimb movements is scored on a scale ranging from 0 to 3. Score 0 is given for a normal forelimb placement. When the forelimb movement is delayed but recovery of postural balance detected, score 1 is given. Score 2 represents a clear, yet insufficient, forelimb reaction, as evidenced by muscle contraction, but lack of success in recovering balance, and score 3 is given for no reaction of movement. The test is repeated three times a day on each side for three consecutive days after an initial training period of three days prior to the first testing.

Staircase Test (Paw Reaching)

A modified version of the staircase test is used for evaluation of paw reaching behavior three weeks following primary and secondary lesion placement. Plexiglass test boxes with a central platform and a removable staircase on each side are used. The apparatus is designed such that only the paw on the same side at each staircase can be used, thus providing a measure of independent forelimb use. For each test the animals are left in the test boxes for 15 min. The double staircase is filled with 7 x 3 chow pellets (Precision food pellets, formula: P, purified rodent diet, size 45 mg; Sandown Scientific) on each side. After each test the number of pellets eaten (successfully retrieved pellets) and the number of pellets taken (touched but dropped) for each paw and the success rate (pellets eaten/pellets taken) are counted separately. After three days of food deprivation (12 g per animal per day) the animals are tested for 11 days. Full analysis is conducted only for the last five days.

MPTP treatment

The neurotoxin 1-methyl-4-phenyl-1,2,3,6-tetrahydro-pyridine (MPTP) causes degeneration of mesencephalic dopaminergic (DAergic) neurons in rodents, non-human primates, and humans and, in so doing, reproduces many of the symptoms of Parkinson's disease. MPTP leads to a marked decrease in the levels of dopamine and its metabolites, and in the number of dopaminergic terminals in the striatum as well as severe loss of the tyrosine hydroxylase (TH)-immunoreactive cell bodies in the substantia nigra, pars compacta.

In order to obtain severe and long-lasting lesions, and to reduce mortality, animals receive single injections of MPTP, and are then tested for severity of lesion 7–10 days later. Successive MPTP injections are administered on days 1, 2 and 3. Animals receive application of 4 mg/kg MPTP hydrochloride (Sigma) in saline once daily. All injections are intraperitoneal (i.p.) and the MPTP stock solution is frozen between injections. Animals are decapitated on day 11.

Immunohistology

At the completion of behavioral experiments, all animals are anaesthetized with 3 ml
5 thiopental (1 g/40 ml i.p., Tyrol Pharma). The mice are perfused transcardially with
0.01 M PBS (pH 7.4) for 2 min, followed by 4% paraformaldehyde (Merck) in PBS
for 15 min. The brains are removed and placed in 4% paraformaldehyde for 24 h at
4 °C. For dehydration they are then transferred to a 20% sucrose (Merck) solution in
0.1 M PBS at 4 °C until they sink. The brains are frozen in methylbutan at -20 °C for
10 2 min and stored at -70 °C. Using a sledge microtome (mod. 3800-Frigocut, Leica),
25 µm sections are taken from the genu of the corpus callosum (AP 1.7 mm) to the
hippocampus (AP 21.8 mm) and from AP 24.16 to AP 26.72. Forty-six sections are
cut and stored in assorters in 0.25 M Tris buffer (pH 7.4) for immunohistochemistry.

15 A series of sections is processed for free-floating tyrosine hydroxylase (TH)
immunohistochemistry. Following three rinses in 0.1 M PBS, endogenous per-
oxidase activity is quenched for 10 min in 0.3% H₂O₂ ±PBS. After rinsing in PBS,
sections are preincubated in 10% normal bovine serum (Sigma) for 5 min as blocking
agent and transferred to either primary anti-rat TH rabbit antiserum (dilution 1:2000).

20

Following overnight incubation at room temperature, sections for TH immuno-
reactivity are rinsed in PBS (2 x10 min) and incubated in biotinylated anti-rabbit
immunoglobulin G raised in goat (dilution 1:200) (Vector) for 90 min, rinsed
repeatedly and transferred to Vectastain ABC (Vector) solution for 1 h. 3,3'
25 -Diaminobenzidine tetrahydrochloride (DAB; Sigma) in 0.1 M PBS, supplemented
with 0.005% H₂O₂ , serves as chromogen in the subsequent visualization reaction.
Sections are mounted on to gelatin-coated slides, left to dry overnight,
counter-stained with hematoxylin dehydrated in ascending alcohol concentrations
and cleared in butylacetate. Coverslips are mounted on entellan.

30

Rotarod Test

We use a modification of the procedure described by Rozas and Labandeira-Garcia (1997), with a CR-1 Rotamex system (Columbus Instruments, Columbus, OH) comprising an IBM-compatible personal computer, a CIO-24 data acquisition card, a control unit, and a four-lane rotarod unit. The rotarod unit consists of a rotating spindle (diameter 7.3 cm) and individual compartments for each mouse. The system software allows preprogramming of session protocols with varying rotational speeds (0–80 rpm). Infrared beams are used to detect when a mouse has fallen onto the base grid beneath the rotarod. The system logs the fall as the end of the experiment for that mouse, and the total time on the rotarod, as well as the time of the fall and all the set-up parameters, are recorded. The system also allows a weak current to be passed through the base grid, to aid training.

3. Dementia

The object recognition task

The object recognition task has been designed to assess the effects of experimental manipulations on the cognitive performance of rodents. A rat is placed in an open field, in which two identical objects are present. The rats inspect both objects during the first trial of the object recognition task. In a second trial, after a retention interval of for example 24 hours, one of the two objects used in the first trial, the ‘familiar’ object, and a novel object are placed in the open field. The inspection time at each of the objects is registered. The basic measures in the OR task is the time spent by a rat exploring the two object the second trial. Good retention is reflected by higher exploration times towards the novel than the ‘familiar’ object.

Administration of the putative cognition enhancer prior to the first trial predominantly allows assessment of the effects on acquisition, and eventually on consolidation processes. Administration of the testing compound after the first trial

allows to assess the effects on consolidation processes, whereas administration before the second trial allows to measure effects on retrieval processes.

The passive avoidance task

5

The passive avoidance task assesses memory performance in rats and mice. The inhibitory avoidance apparatus consists of a two-compartment box with a light compartment and a dark compartment. The two compartments are separated by a guillotine door that can be operated by the experimenter. A threshold of 2 cm
10 separates the two compartments when the guillotine door is raised. When the door is open, the illumination in the dark compartment is about 2 lux. The light intensity is about 500 lux at the center of the floor of the light compartment.

Two habituation sessions, one shock session, and a retention session are given,
15 separated by inter-session intervals of 24 hours. In the habituation sessions and the retention session the rat is allowed to explore the apparatus for 300 sec. The rat is placed in the light compartment, facing the wall opposite to the guillotine door. After an accommodation period of 15 sec. the guillotine door is opened so that all parts of the apparatus can be visited freely. Rats normally avoid brightly lit areas and will
20 enter the dark compartment within a few seconds.

In the shock session the guillotine door between the compartments is lowered as soon as the rat has entered the dark compartment with its four paws, and a scrambled 1 mA footshock is administered for 2 sec. The rat is removed from the apparatus and
25 put back into its home cage. The procedure during the retention session is identical to that of the habituation sessions.

The step-through latency, that is the first latency of entering the dark compartment (in sec.) during the retention session is an index of the memory performance of the
30 animal; the longer the latency to enter the dark compartment, the better the retention is. A testing compound is given half an hour before the shock session, together with

- 77 -

1 mg*kg⁻¹ scopolamine. Scopolamine impairs the memory performance during the retention session 24 hours later. If the test compound increases the enter latency compared with the scopolamine-treated controls, is likely to possess cognition enhancing potential.

5

The Morris water escape task

The Morris water escape task measures spatial orientation learning in rodents. It is a test system that has extensively been used to investigate the effects of putative therapeutic on the cognitive functions of rats and mice. The performance of an animal is assessed in a circular water tank with an escape platform that is submerged about 1 cm below the surface of the water. The escape platform is not visible for an animal swimming in the water tank. Abundant extra-maze cues are provided by the furniture in the room, including desks, computer equipment, a second water tank, the presence of the experimenter, and by a radio on a shelf that is playing softly.

15

The animals receive four trials during five daily acquisition sessions. A trial is started by placing an animal into the pool, facing the wall of the tank. Each of four starting positions in the quadrants north, east, south, and west is used once in a series of four trials; their order is randomized. The escape platform is always in the same position. A trial is terminated as soon as the animal had climbs onto the escape platform or when 90 seconds have elapsed, whichever event occurs first. The animal is allowed to stay on the platform for 30 seconds. Then it is taken from the platform and the next trial is started. If an animal did not find the platform within 90 seconds it is put on the platform by the experimenter and is allowed to stay there for 30 seconds. After the fourth trial of the fifth daily session, an additional trial is given as a probe trial: the platform is removed, and the time the animal spends in the four quadrants is measured for 30 or 60 seconds. In the probe trial, all animals start from the same start position, opposite to the quadrant where the escape platform had been positioned during acquisition.

20

25

30

- Four different measures are taken to evaluate the performance of an animal during acquisition training: escape latency, traveled distance, distance to platform, and swimming speed. The following measures are evaluated for the probe trial: time (s) in quadrants and traveled distance (cm) in the four quadrants. The probe trial provides additional information about how well an animal learned the position of the escape platform. If an animal spends more time and swims a longer distance in the quadrant where the platform had been positioned during the acquisition sessions than in any other quadrant, one concludes that the platform position has been learned well.
- 5
- 10 In order to assess the effects of putative cognition enhancing compounds, rats or mice with specific brain lesions which impair cognitive functions, or animals treated with compounds such as scopolamine or MK-801, which interfere with normal learning, or aged animals which suffer from cognitive deficits, are used.

The T-maze spontaneous alternation task

The T-maze spontaneous alternation task (TeMCAT) assesses the spatial memory performance in mice. The start arm and the two goal arms of the T-maze are provided with guillotine doors which can be operated manually by the experimenter. A mouse is put into the start arm at the beginning of training. The guillotine door is closed. In the first trial, the 'forced trial', either the left or right goal arm is blocked by lowering the guillotine door. After the mouse has been released from the start arm, it will negotiate the maze, eventually enter the open goal arm, and return to the start position, where it will be confined for 5 seconds, by lowering the guillotine door. Then, the animal can choose freely between the left and right goal arm (all guillotine-doors opened) during 14 'free choice' trials. As soon as the mouse has entered one goal arm, the other one is closed. The mouse eventually returns to the start arm and is free to visit whichever goal arm it wants after having been confined to the start arm for 5 seconds. After completion of 14 free choice trials in one session, the animal is removed from the maze. During training, the animal is never handled.

The percent alternations out of 14 trials is calculated. This percentage and the total time needed to complete the first forced trial and the subsequent 14 free choice trials (in s) is analyzed. Cognitive deficits are usually induced by an injection of scopolamine, 30 min before the start of the training session. Scopolamine reduced the per-cent alternations to chance level, or below. A cognition enhancer, which is always administered before the training session, will at least partially, antagonize the scopolamine-induced reduction in the spontaneous alternation rate.

REFERENCES

1: Di Cunto F, Imarisio S, Hirsch E, Broccoli V, Bulfone A, Migheli A, Atzori C, Turco E, Triolo R, Dotto GP, Silengo L, Altruda F. Defective

neurogenesis in citron kinase knockout mice by altered cytokinesis and massive apoptosis. *Neuron*. 2000 Oct;28(1):115-27.

2: Di Cunto F, Calautti E, Hsiao J, Ong L, Topley G, Turco E, Dotto GP. Citron rho-interacting kinase, a novel tissue-specific ser/thr kinase encompassing the
5 Rho-Rac-binding protein Citron. *J Biol Chem*. 1998 Nov 6;273(45):29706-11.

3: Madaule P, Furuyashiki T, Reid T, Ishizaki T, Watanabe G, Morii N, Narumiya S. A novel partner for the GTP-bound forms of rho and rac. *FEBS Lett*. 1995 Dec 18;377(2):243-8.

4: Fujisawa K, Madaule P, Ishizaki T, Watanabe G, Bito H, Saito Y, Hall A,
10 Narumiya S. Different regions of Rho determine Rho-selective binding of different classes of Rho target molecules. *J Biol Chem*. 1998 Jul 24;273(30):18943-9.

CLAIMS

1. An isolated polynucleotide being selected from the group consisting of:
 - a. a polynucleotide encoding a human citron rho/rac-interacting kinase polypeptide comprising an amino acid sequence selected from the group consisting of:
 - i. amino acid sequences which are at least about 97% identical to the amino acid sequence shown in SEQ ID NO: 2; and
 - ii. the amino acid sequence shown in SEQ ID NO: 2.
 - b. a polynucleotide comprising the sequence of SEQ ID NOS: 1 or 24;
 - c. a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) and (b) and encodes a human citron rho/rac-interacting kinase polypeptide;
 - d. a polynucleotide the sequence of which deviates from the polynucleotide sequences specified in (a) to (c) due to the degeneration of the genetic code and encodes a human citron rho/rac-interacting kinase polypeptide; and
 - e. a polynucleotide which represents a fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (d) and encodes a human citron rho/rac-interacting kinase polypeptide.
2. An expression vector containing any polynucleotide of claim 1.
3. A host cell containing the expression vector of claim 2.
4. A substantially purified human citron rho/rac-interacting kinase polypeptide encoded by a polynucleotide of claim 1.
5. A method for producing a human citron rho/rac-interacting kinase polypeptide, wherein the method comprises the following steps:

- a. culturing the host cell of claim 3 under conditions suitable for the expression of the human citron rho/rac-interacting kinase polypeptide; and
 - b. recovering the human citron rho/rac-interacting kinase polypeptide from the host cell culture.
- 5
6. A method for detection of a polynucleotide encoding a human citron rho/rac-interacting kinase polypeptide in a biological sample comprising the following steps:
 - 10 a. hybridizing any polynucleotide of claim 1 to a nucleic acid material of a biological sample, thereby forming a hybridization complex; and
 - b. detecting said hybridization complex.
7. The method of claim 6, wherein before hybridization, the nucleic acid material of the biological sample is amplified.
- 15
8. A method for the detection of a polynucleotide of claim 1 or a human citron rho/rac-interacting kinase polypeptide of claim 4 comprising the steps of:
 - 20 a. contacting a biological sample with a reagent which specifically interacts with the polynucleotide or the human citron rho/rac-interacting kinase polypeptide and
 - b. detecting the interaction
9. A diagnostic kit for conducting the method of any one of claims 6 to 8.
- 25
10. A method of screening for agents which decrease the activity of a human citron rho/rac-interacting kinase, comprising the steps of:
 - a. contacting a test compound with any human citron rho/rac-interacting kinase polypeptide encoded by any polynucleotide of claim 1;
 - 30 b. detecting binding of the test compound to the human citron rho/rac-interacting kinase polypeptide, wherein a test compound which binds

to the polypeptide is identified as a potential therapeutic agent for decreasing the activity of a human citron rho/rac-interacting kinase.

- 5 11. A method of screening for agents which regulate the activity of a human citron rho/rac-interacting kinase, comprising the steps of:
- a. contacting a test compound with a human citron rho/rac-interacting kinase polypeptide encoded by any polynucleotide of claim 1; and
 - b. detecting a human citron rho/rac-interacting kinase activity of the polypeptide, wherein a test compound which increases the human
10 citron rho/rac-interacting kinase activity is identified as a potential therapeutic agent for increasing the activity of the human citron rho/rac-interacting kinase, and wherein a test compound which decreases the human citron rho/rac-interacting kinase activity of the polypeptide is identified as a potential therapeutic agent for decreasing
15 the activity of the human citron rho/rac-interacting kinase.
12. A method of screening for agents which decrease the activity of a human citron rho/rac-interacting kinase, comprising the steps of:
- a. contacting a test compound with any polynucleotide of claim 1 and
20 detecting binding of the test compound to the polynucleotide, wherein a test compound which binds to the polynucleotide is identified as a potential therapeutic agent for decreasing the activity of human citron rho/rac-interacting kinase.
- 25 13. A method of reducing the activity of human citron rho/rac-interacting kinase, comprising the steps of:
- a. contacting a cell with a reagent which specifically binds to any polynucleotide of claim 1 or any human citron rho/rac-interacting kinase polypeptide of claim 4, whereby the activity of human citron
30 rho/rac-interacting kinase is reduced.

14. A reagent that modulates the activity of a human citron rho/rac-interacting kinase polypeptide or a polynucleotide wherein said reagent is identified by the method of any of the claim 10 to 12.
- 5 15. A pharmaceutical composition, comprising:
 - a. the expression vector of claim 2 or the reagent of claim 14 and a pharmaceutically acceptable carrier.
- 10 16. Use of the expression vector of claim 2 or the reagent of claim 14 in the preparation of a medicament for modulating the activity of a human citron rho/rac-interacting kinase in a disease.
17. Use of claim 16 wherein the disease is obesity, a CNS disorder or COPD.

- 1/84 -

Fig. 1

atgttgaagt	tcaaatatgg	agcgcggaat	cctttggatg	ctggtgctgc	tgaaccatt	60
gccagccggg	cctccagggt	gaatctgttc	ttccagggga	aaccaccctt	tatgactcaa	120
cagcagatgt	ctcctctttc	ccgagaaggg	atattagatg	ccctctttgt	tctctttgaa	180
gaatgcagtc	agcctgctct	gatgaagatt	aagcacgtga	gcaactttgt	ccggaagtat	240
tccgacacca	tagctgagtt	acaggagctc	cagccttcgg	caaaggactt	cgaagtcaga	300
agtctttag	gttggtgca	ctttgctgaa	gtgcagggtg	taagagagaa	agcaaccggg	360
gacatctatg	ctatgaaagt	gatgaagaag	aaggctttat	tggcccagga	gcaggtttca	420
ttttttgagg	aagagcggaa	catattatct	cgaagcacia	gcccgtggat	cccccaatta	480
cagtatgcct	ttcaggacaa	aatcacctt	tatctggtca	tggaaatca	gcctggaggg	540
gacttgctgt	cacttttgaa	tagatatgag	gaccagttag	atgaaaacct	gatacagttt	600
tacctagctg	agctgatttt	ggctgttcac	agcgttcac	tgatgggata	cgtgcatcga	660
gacatcaagc	ctgagaacat	tctcgttgac	cgcacaggac	acatcaagct	gggtggatttt	720
ggatctgccg	cgaaaatgaa	ttcaaacaag	atggtgaatg	ccaaactccc	gattgggacc	780
ccagattaca	tggctcctga	agtgtgact	gtgatgaacg	gggatggaaa	aggcacctac	840
ggcctggact	gtgactggtg	gtcagtgggc	gtgattgcct	atgagatgat	ttatgggaga	900
ttccccttcg	cagagggaa	ctctgccaga	accttcaata	acattatgaa	tttccagcgg	960
tttttgaaat	ttccagatga	ccccaaagt	agcagtgact	ttcttgatct	gattcaaagc	1020
ttgttgctgc	gccagaaaga	gagactgaag	tttgaaggct	tttgctgcca	tcctttcttc	1080
tctaaaattg	actggaacaa	cattcgtaac	tctcctcccc	ccttcgttcc	cacctcaag	1140
tctgacgatg	acacctccaa	ttttgatgaa	ccagagaaga	attcgtgggt	ttcatcctct	1200
ccgtgccagc	tgagcccctc	aggcttctcg	gggtgaagaac	tgccgtttgt	ggggttttcg	1260
tacagcaagg	cactggggat	tcttggtaga	tctgagtctg	ttgtgtcggg	tctggactcc	1320
cctgccaaaga	ctagctccat	ggaaaagaaa	cttctcatca	aaagcaaaga	gctacaagac	1380
tctcaggaca	agtgtcacia	gatggagcag	gaaatgacct	ggttacatcg	gagagtgtca	1440
gaggtggagg	ctgtgcttag	tcagaaggag	gtggagctga	aggcctctga	gactcagaga	1500
tcctcctcgg	agcaggacct	tgctacctac	atcacagaat	gcagtagctt	aaagcgaagt	1560
ttggagcaag	cacggatgga	ggtgtcccag	gaggatgaca	aagcactgca	gcttctccat	1620
gatatcagag	agcagagccg	gaagctccaa	gaaatcaaag	agcaggagta	ccaggctcaa	1680
gtggaagaaa	tgaggttgat	gatgaatcag	ttggaagagg	atcttgtctc	agcaagaaga	1740
cggagtgatc	tctacgaatc	tgagctgaga	gagtctcggc	ttgctgctga	agaattcaag	1800
cggaaagcga	cagaatgtca	gcataaactg	ttgaaggcta	aggatcaagg	gaagcctgaa	1860
gtgggagaat	atgcgaaact	ggagaagatc	aatgctgagc	agcagctcaa	aattcaggag	1920
ctccaagaga	aactggagaa	ggctgtaaaa	gccagcacgg	aggccaccga	gctgctgcag	1980
aatatccgcc	aggcaaagga	gcgagccgag	aggagctgg	agaagctgca	gaaccgagag	2040
gattcttctg	aaggcatcag	aaagaagctg	gtggaagctg	aggaacgccg	ccattctctg	2100
gagaacaagg	taaagagact	agagaccatg	gagcgtagag	aaaacagact	gaaggatgac	2160
atccagacaa	aatcccaaca	gatccagcag	atggctgata	aaattctgga	gctcgaagag	2220
aaacatcggg	aggcccaagt	ctcagcccag	cacctagaag	tgcacctgaa	acagaaagag	2280

- 2/84 -

Fig. 1 (continued)

cagcactatg	aggaaaagat	taaagtgttg	gacaatcaga	taaagaaaga	cctggctgac	2340
aaggagacac	tggagaacat	gatgcagaga	cacgaggagg	aggcccatga	gaagggcaaa	2400
attctcagcg	aacagaaggc	gatgatcaat	gctatggatt	ccaagatcag	atccctggaa	2460
cagaggattg	tggaaactgtc	tgaagccaat	aaacttgacg	caaatagcag	tctttttacc	2520
caaaggaaca	tgaaggccca	agaagagatg	atttctgaac	tcaggcaaca	gaaattttac	2580
ctggagacac	aggctgggaa	gttggaggcc	cagaaccgaa	aactggagga	gcagctggag	2640
aagatcagcc	accaagacca	cagtgcacaag	aatcggtgcg	tggaaactgga	gacaagattg	2700
cgggagggtca	gtctagagca	cgaggagcag	aaactggagc	tcaagcgcca	gctcacagag	2760
ctacagctct	ccctgcagga	gcgcgagtcg	cagttgacag	ccctgcaggc	tgcacgggcg	2820
gccctggaga	gccagcttcg	ccaggcgaag	acagagctgg	aagagaccac	agcagaagct	2880
gaagaggaga	tccaggcact	cacggcacat	agagatgaaa	tccagcgcaa	atttgatgct	2940
cttcgtaaca	gctgtactgt	aatcacagac	ctggaggagc	agctaaacca	gctgaccgag	3000
gacaacgctg	aactcaacaa	ccaaaacttc	tacttgtcca	aacaactcga	tgaggcttct	3060
ggcgccaacg	acgagattgt	acaactgcga	agtgaagtgg	accatctccg	ccgggagatc	3120
acggaacgag	agatgcagct	taccagccag	aagcaaacga	tggaggctct	gaagaccacg	3180
tgcaccatgc	tggaggaaca	ggtcattgat	ttggaggccc	taaacgatga	gctgctagaa	3240
aaagagcggc	agtgggaggc	ctggaggagc	gtcctgggtg	atgagaaatc	ccagtttgag	3300
tgtcgggttc	gagagctgca	gagaatgctg	gacaccgaga	aacagagcag	ggcgagagcc	3360
gatcagcgga	tcaccgagtc	tgcgcagggtg	gtggagctgg	cagtgaagga	gcacaaggct	3420
gagattctcg	ctctgcagca	ggctctcaaa	gagcagaagc	tgaaggccga	gagcctctct	3480
gacaagctca	atgacctgga	gaagaagcat	gctatgcttg	aaatgaatgc	ccgaagctta	3540
cagcagaagc	tggagactga	acgagagctc	aaacagaggc	ttctggaaga	gcaagccaaa	3600
ttacagcagc	agatggacct	gcagaaaaat	cacattttcc	gtctgactca	aggactgcaa	3660
gaagctctag	atcgggctga	tctactgaag	acagaaagaa	gtgacttgga	gtatcagctg	3720
gaaaacattc	aggttctcta	ttctcatgaa	aagggtgaaa	tggaaaggcac	tattttctcaa	3780
caaaccaaac	tcatgtattt	tctgcaagcc	aaaatggacc	aacctgctaa	aaagaaaaag	3840
gttcctctgc	agtacaatga	gctgaagctg	gccctggaga	aggagaaagc	tcgctgtgca	3900
gagctagagg	aagcccttca	gaagaccgcg	atcgagctcc	ggtccgcccg	ggaggaagct	3960
gccaccgcga	aagcaacgga	ccaccacac	ccatccacgc	cagccaccgc	gaggcagcag	4020
atcgccatgt	ccgccatcgt	gcggtcgcca	gagcaccagc	ccagtgccat	gagcctgctg	4080
gcccgcgcat	ccagccgcag	aaaggagtct	tcaactccag	aggaatttag	tcggcgctct	4140
aaggaacgca	tgcaccacaa	tattcctcac	cgattcaacg	taggactgaa	catgagagcc	4200
acaaagtgtg	ctgtgtgtct	ggataccgtg	cactttggac	gccaggcatc	caaattgtctc	4260
gaatgtcagg	tgatgtgtca	ccccaaagtc	tccacgtgct	tgccagccac	ctgaggcttg	4320
cctgctgaat	atgccacaca	cttcaccgag	gccttctgcc	gtgacaaaat	gaactcccca	4380
ggtctccaga	ccaaggagcc	cagcagcagc	ttgcacctgg	aaggggtggat	gaaggtgccc	4440
aggaataaca	aacgaggaca	gcaaggctgg	gacaggaagt	acattgtcct	ggagggatca	4500
aaagtcctca	tttatgacaa	tgaagccaga	gaagctggac	agaggccggt	ggaagaattt	4560

- 3/84 -

Fig. 1 (continued)

gagctgtgcc	ttcccgaacg	ggatgtatct	attcatggtg	ccgttggtgc	ttccgaactc	4620
gcaaatacag	ccaaagcaga	tgtcccatac	atactgaaga	tggaaatctca	cccgcacacc	4680
acctgctggc	ccgggagaa	cctctacttg	ctagctccca	gcttccctga	caaacagcgc	4740
tgggtcaccg	ccttagaatc	agttgtcgca	ggtgggagag	tttctagggg	aaaagcagaa	4800
gctgatgcta	aactgcttgg	aaactccctg	ctgaaactgg	aaggtgatga	ccgtctagac	4860
atgaactgca	cgctgccctt	cagtgaccag	gtggtgttgg	tgggcaccga	ggaagggctc	4920
tacgccctga	atgtcttgaa	aaactcccta	acccatgtcc	caggaattgg	agcagtcttc	4980
caaatttata	ttatcaagga	cctggagaag	ctactcatga	tagcaggaga	agagcgggca	5040
ctgtgtcttg	tggacgtgaa	gaaagtgaag	cagtccctgg	cccagtccca	cctgcctgcc	5100
cagcccgaca	tctcacccaa	cattttttgaa	gctgtcaagg	gctgccactt	gtttggggca	5160
ggcaagattg	agaacgggct	ctgcatctgt	gcagccatgc	ccagcaaagt	cgctattctc	5220
cgctacaacg	aaaacctcag	caaatactgc	atccggaaag	agatagagac	ctcagagccc	5280
tgcagctgta	tccacttcac	caattacagt	atcctcattg	gaaccaataa	attctacgaa	5340
atcgacatga	agcagtacac	gctcgaggaa	ttcctggata	agaatgacca	ttccttgga	5400
cctgctgtgt	ttgcgcctc	ttccaacagc	ttcctgtct	caatcgtgca	ggtgaacagc	5460
gcagggcagc	gagaggagta	cttgctgtgt	ttcacgaat	ttggagtgtt	cgtaggattct	5520
tacggaagac	gtagccgcac	agacgatctc	aagtggagtc	gcttaccttt	ggcctttgcc	5580
tacagagaac	cctatctgtt	tgtgaccac	ttcaactcac	tcgaagtaat	tgagatccag	5640
gcacgctcct	cagcagggac	ccctgcccga	gcgtacctgg	acatcccga	cccgcgctac	5700
ctgggccctg	ccatttcctc	aggagcgatt	tacttggcgt	cctcatacca	ggataaatta	5760
agggtcattt	gctgcaagg	aaacctcgtg	aaggagtccg	gcactgaaca	ccaccggggc	5820
ccgtccacct	cccgcagcag	ccccacaag	cgaggccac	ccacgtacaa	cgagcacatc	5880
accaagcgcg	tggcctccag	cccagcgccg	cccgaaggcc	ccagccacc	gagagagcca	5940
agcacacccc	accgctaccg	cgagggcgcg	accgagctgc	gcagggacaa	gtctcctggc	6000
cgccccctgg	agcgagagaa	gtcccccggc	cggatgctca	gcacgcggag	agagcggtcc	6060
ccggggaggc	tgtttgaaga	cagcagcagg	ggccggctgc	ctgcgggagc	cgtaggagcc	6120
ccgctgtccc	aggtgaacaa	ggtctgggac	cagtcttcag	tataa		6165

- 4/84 -

Fig. 2

Met Leu Lys Phe Lys Tyr Gly Ala Arg Asn Pro Leu Asp Ala Gly Ala
 1 5 10 15
 Ala Glu Pro Ile Ala Ser Arg Ala Ser Arg Leu Asn Leu Phe Phe Gln
 20 25 30
 Gly Lys Pro Pro Phe Met Thr Gln Gln Gln Met Ser Pro Leu Ser Arg
 35 40 45
 Glu Gly Ile Leu Asp Ala Leu Phe Val Leu Phe Glu Glu Cys Ser Gln
 50 55 60
 Pro Ala Leu Met Lys Ile Lys His Val Ser Asn Phe Val Arg Lys Tyr
 65 70 75 80
 Ser Asp Thr Ile Ala Glu Leu Gln Glu Leu Gln Pro Ser Ala Lys Asp
 85 90 95
 Phe Glu Val Arg Ser Leu Val Gly Cys Gly His Phe Ala Glu Val Gln
 100 105 110
 Val Val Arg Glu Lys Ala Thr Gly Asp Ile Tyr Ala Met Lys Val Met
 115 120 125
 Lys Lys Lys Ala Leu Leu Ala Gln Glu Gln Val Ser Phe Phe Glu Glu
 130 135 140
 Glu Arg Asn Ile Leu Ser Arg Ser Thr Ser Pro Trp Ile Pro Gln Leu
 145 150 155 160
 Gln Tyr Ala Phe Gln Asp Lys Asn His Leu Tyr Leu Val Met Glu Tyr
 165 170 175
 Gln Pro Gly Gly Asp Leu Leu Ser Leu Leu Asn Arg Tyr Glu Asp Gln
 180 185 190
 Leu Asp Glu Asn Leu Ile Gln Phe Tyr Leu Ala Glu Leu Ile Leu Ala
 195 200 205
 Val His Ser Val His Leu Met Gly Tyr Val His Arg Asp Ile Lys Pro
 210 215 220
 Glu Asn Ile Leu Val Asp Arg Thr Gly His Ile Lys Leu Val Asp Phe
 225 230 235 240
 Gly Ser Ala Ala Lys Met Asn Ser Asn Lys Met Val Asn Ala Lys Leu
 245 250 255
 Pro Ile Gly Thr Pro Asp Tyr Met Ala Pro Glu Val Leu Thr Val Met
 260 265 270
 Asn Gly Asp Gly Lys Gly Thr Tyr Gly Leu Asp Cys Asp Trp Trp Ser
 275 280 285
 Val Gly Val Ile Ala Tyr Glu Met Ile Tyr Gly Arg Ser Pro Phe Ala
 290 295 300

- 5/84 -

Fig. 2 (continued)

Glu Gly Thr Ser Ala Arg Thr Phe Asn Asn Ile Met Asn Phe Gln Arg			
305	310	315	320
Phe Leu Lys Phe Pro Asp Asp Pro Lys Val Ser Ser Asp Phe Leu Asp			
	325	330	335
Leu Ile Gln Ser Leu Leu Cys Gly Gln Lys Glu Arg Leu Lys Phe Glu			
	340	345	350
Gly Leu Cys Cys His Pro Phe Phe Ser Lys Ile Asp Trp Asn Asn Ile			
	355	360	365
Arg Asn Ser Pro Pro Pro Phe Val Pro Thr Leu Lys Ser Asp Asp Asp			
	370	375	380
Thr Ser Asn Phe Asp Glu Pro Glu Lys Asn Ser Trp Val Ser Ser Ser			
385	390	395	400
Pro Cys Gln Leu Ser Pro Ser Gly Phe Ser Gly Glu Glu Leu Pro Phe			
	405	410	415
Val Gly Phe Ser Tyr Ser Lys Ala Leu Gly Ile Leu Gly Arg Ser Glu			
	420	425	430
Ser Val Val Ser Gly Leu Asp Ser Pro Ala Lys Thr Ser Ser Met Glu			
	435	440	445
Lys Lys Leu Leu Ile Lys Ser Lys Glu Leu Gln Asp Ser Gln Asp Lys			
	450	455	460
Cys His Lys Met Glu Gln Glu Met Thr Arg Leu His Arg Arg Val Ser			
465	470	475	480
Glu Val Glu Ala Val Leu Ser Gln Lys Glu Val Glu Leu Lys Ala Ser			
	485	490	495
Glu Thr Gln Arg Ser Leu Leu Glu Gln Asp Leu Ala Thr Tyr Ile Thr			
	500	505	510
Glu Cys Ser Ser Leu Lys Arg Ser Leu Glu Gln Ala Arg Met Glu Val			
	515	520	525
Ser Gln Glu Asp Asp Lys Ala Leu Gln Leu Leu His Asp Ile Arg Glu			
	530	535	540
Gln Ser Arg Lys Leu Gln Glu Ile Lys Glu Gln Glu Tyr Gln Ala Gln			
545	550	555	560
Val Glu Glu Met Arg Leu Met Met Asn Gln Leu Glu Glu Asp Leu Val			
	565	570	575
Ser Ala Arg Arg Arg Ser Asp Leu Tyr Glu Ser Glu Leu Arg Glu Ser			
	580	585	590
Arg Leu Ala Ala Glu Glu Phe Lys Arg Lys Ala Thr Glu Cys Gln His			
	595	600	605

- 6/84 -

Fig. 2 (continued)

Lys	Leu	Leu	Lys	Ala	Lys	Asp	Gln	Gly	Lys	Pro	Glu	Val	Gly	Glu	Tyr
610					615					620					
Ala	Lys	Leu	Glu	Lys	Ile	Asn	Ala	Glu	Gln	Gln	Leu	Lys	Ile	Gln	Glu
625				630					635					640	
Leu	Gln	Glu	Lys	Leu	Glu	Lys	Ala	Val	Lys	Ala	Ser	Thr	Glu	Ala	Thr
			645					650					655		
Glu	Leu	Leu	Gln	Asn	Ile	Arg	Gln	Ala	Lys	Glu	Arg	Ala	Glu	Arg	Glu
			660				665						670		
Leu	Glu	Lys	Leu	Gln	Asn	Arg	Glu	Asp	Ser	Ser	Glu	Gly	Ile	Arg	Lys
	675					680					685				
Lys	Leu	Val	Glu	Ala	Glu	Glu	Arg	Arg	His	Ser	Leu	Glu	Asn	Lys	Val
	690				695						700				
Lys	Arg	Leu	Glu	Thr	Met	Glu	Arg	Arg	Glu	Asn	Arg	Leu	Lys	Asp	Asp
705				710					715					720	
Ile	Gln	Thr	Lys	Ser	Gln	Gln	Ile	Gln	Gln	Met	Ala	Asp	Lys	Ile	Leu
			725				730						735		
Glu	Leu	Glu	Glu	Lys	His	Arg	Glu	Ala	Gln	Val	Ser	Ala	Gln	His	Leu
			740				745						750		
Glu	Val	His	Leu	Lys	Gln	Lys	Glu	Gln	His	Tyr	Glu	Glu	Lys	Ile	Lys
	755					760						765			
Val	Leu	Asp	Asn	Gln	Ile	Lys	Lys	Asp	Leu	Ala	Asp	Lys	Glu	Thr	Leu
	770					775						780			
Glu	Asn	Met	Met	Gln	Arg	His	Glu	Glu	Glu	Ala	His	Glu	Lys	Gly	Lys
785				790					795					800	
Ile	Leu	Ser	Glu	Gln	Lys	Ala	Met	Ile	Asn	Ala	Met	Asp	Ser	Lys	Ile
			805					810					815		
Arg	Ser	Leu	Glu	Gln	Arg	Ile	Val	Glu	Leu	Ser	Glu	Ala	Asn	Lys	Leu
		820					825						830		
Ala	Ala	Asn	Ser	Ser	Leu	Phe	Thr	Gln	Arg	Asn	Met	Lys	Ala	Gln	Glu
	835					840						845			
Glu	Met	Ile	Ser	Glu	Leu	Arg	Gln	Gln	Lys	Phe	Tyr	Leu	Glu	Thr	Gln
	850					855					860				
Ala	Gly	Lys	Leu	Glu	Ala	Gln	Asn	Arg	Lys	Leu	Glu	Glu	Gln	Leu	Glu
865				870					875					880	
Lys	Ile	Ser	His	Gln	Asp	His	Ser	Asp	Lys	Asn	Arg	Leu	Leu	Glu	Leu
			885					890						895	

- 7/84 -

Fig. 2 (continued)

Glu Thr Arg Leu Arg Glu Val Ser Leu Glu His Glu Glu Gln Lys Leu			
900	905	910	
Glu Leu Lys Arg Gln Leu Thr Glu Leu Gln Leu Ser Leu Gln Glu Arg			
915	920	925	
Glu Ser Gln Leu Thr Ala Leu Gln Ala Ala Arg Ala Ala Leu Glu Ser			
930	935	940	
Gln Leu Arg Gln Ala Lys Thr Glu Leu Glu Glu Thr Thr Ala Glu Ala			
945	950	955	960
Glu Glu Glu Ile Gln Ala Leu Thr Ala His Arg Asp Glu Ile Gln Arg			
965	970	975	
Lys Phe Asp Ala Leu Arg Asn Ser Cys Thr Val Ile Thr Asp Leu Glu			
980	985	990	
Glu Gln Leu Asn Gln Leu Thr Glu Asp Asn Ala Glu Leu Asn Asn Gln			
995	1000	1005	
Asn Phe Tyr Leu Ser Lys Gln Leu Asp Glu Ala Ser Gly Ala Asn Asp			
1010	1015	1020	
Glu Ile Val Gln Leu Arg Ser Glu Val Asp His Leu Arg Arg Glu Ile			
1025	1030	1035	1040
Thr Glu Arg Glu Met Gln Leu Thr Ser Gln Lys Gln Thr Met Glu Ala			
1045	1050	1055	
Leu Lys Thr Thr Cys Thr Met Leu Glu Glu Gln Val Met Asp Leu Glu			
1060	1065	1070	
Ala Leu Asn Asp Glu Leu Leu Glu Lys Glu Arg Gln Trp Glu Ala Trp			
1075	1080	1085	
Arg Ser Val Leu Gly Asp Glu Lys Ser Gln Phe Glu Cys Arg Val Arg			
1090	1095	1100	
Glu Leu Gln Arg Met Leu Asp Thr Glu Lys Gln Ser Arg Ala Arg Ala			
1105	1110	1115	1120
Asp Gln Arg Ile Thr Glu Ser Arg Gln Val Val Glu Leu Ala Val Lys			
1125	1130	1135	
Glu His Lys Ala Glu Ile Leu Ala Leu Gln Gln Ala Leu Lys Glu Gln			
1140	1145	1150	
Lys Leu Lys Ala Glu Ser Leu Ser Asp Lys Leu Asn Asp Leu Glu Lys			
1155	1160	1165	
Lys His Ala Met Leu Glu Met Asn Ala Arg Ser Leu Gln Gln Lys Leu			
1170	1175	1180	
Glu Thr Glu Arg Glu Leu Lys Gln Arg Leu Leu Glu Glu Gln Ala Lys			
1185	1190	1195	1200

- 8/84 -

Fig. 2 (continued)

Leu Gln Gln Gln Met Asp Leu Gln Lys Asn His Ile Phe Arg Leu Thr			
	1205	1210	1215
Gln Gly Leu Gln Glu Ala Leu Asp Arg Ala Asp Leu Leu Lys Thr Glu			
	1220	1225	1230
Arg Ser Asp Leu Glu Tyr Gln Leu Glu Asn Ile Gln Val Leu Tyr Ser			
	1235	1240	1245
His Glu Lys Val Lys Met Glu Gly Thr Ile Ser Gln Gln Thr Lys Leu			
	1250	1255	1260
Ile Asp Phe Leu Gln Ala Lys Met Asp Gln Pro Ala Lys Lys Lys Lys			
	1265	1270	1275 1280
Val Pro Leu Gln Tyr Asn Glu Leu Lys Leu Ala Leu Glu Lys Glu Lys			
	1285	1290	1295
Ala Arg Cys Ala Glu Leu Glu Glu Ala Leu Gln Lys Thr Arg Ile Glu			
	1300	1305	1310
Leu Arg Ser Ala Arg Glu Glu Ala Ala His Arg Lys Ala Thr Asp His			
	1315	1320	1325
Pro His Pro Ser Thr Pro Ala Thr Ala Arg Gln Gln Ile Ala Met Ser			
	1330	1335	1340
Ala Ile Val Arg Ser Pro Glu His Gln Pro Ser Ala Met Ser Leu Leu			
	1345	1350	1355 1360
Ala Pro Pro Ser Ser Arg Arg Lys Glu Ser Ser Thr Pro Glu Glu Phe			
	1365	1370	1375
Ser Arg Arg Leu Lys Glu Arg Met His His Asn Ile Pro His Arg Phe			
	1380	1385	1390
Asn Val Gly Leu Asn Met Arg Ala Thr Lys Cys Ala Val Cys Leu Asp			
	1395	1400	1405
Thr Val His Phe Gly Arg Gln Ala Ser Lys Cys Leu Glu Cys Gln Val			
	1410	1415	1420
Met Cys His Pro Lys Cys Ser Thr Cys Leu Pro Ala Thr Cys Gly Leu			
	1425	1430	1435 1440
Pro Ala Glu Tyr Ala Thr His Phe Thr Glu Ala Phe Cys Arg Asp Lys			
	1445	1450	1455
Met Asn Ser Pro Gly Leu Gln Thr Lys Glu Pro Ser Ser Ser Leu His			
	1460	1465	1470
Leu Glu Gly Trp Met Lys Val Pro Arg Asn Asn Lys Arg Gly Gln Gln			
	1475	1480	1485

- 9/84 -

Fig. 2 (continued)

Gly Trp Asp Arg Lys Tyr Ile Val Leu Glu Gly Ser Lys Val Leu Ile			
1490	1495	1500	
Tyr Asp Asn Glu Ala Arg Glu Ala Gly Gln Arg Pro Val Glu Glu Phe			
1505	1510	1515	1520
Glu Leu Cys Leu Pro Asp Gly Asp Val Ser Ile His Gly Ala Val Gly			
	1525	1530	1535
Ala Ser Glu Leu Ala Asn Thr Ala Lys Ala Asp Val Pro Tyr Ile Leu			
	1540	1545	1550
Lys Met Glu Ser His Pro His Thr Thr Cys Trp Pro Gly Arg Thr Leu			
	1555	1560	1565
Tyr Leu Leu Ala Pro Ser Phe Pro Asp Lys Gln Arg Trp Val Thr Ala			
	1570	1575	1580
Leu Glu Ser Val Val Ala Gly Gly Arg Val Ser Arg Glu Lys Ala Glu			
1585	1590	1595	1600
Ala Asp Ala Lys Leu Leu Gly Asn Ser Leu Leu Lys Leu Glu Gly Asp			
	1605	1610	1615
Asp Arg Leu Asp Met Asn Cys Thr Leu Pro Phe Ser Asp Gln Val Val			
	1620	1625	1630
Leu Val Gly Thr Glu Glu Gly Leu Tyr Ala Leu Asn Val Leu Lys Asn			
	1635	1640	1645
Ser Leu Thr His Val Pro Gly Ile Gly Ala Val Phe Gln Ile Tyr Ile			
	1650	1655	1660
Ile Lys Asp Leu Glu Lys Leu Leu Met Ile Ala Gly Glu Glu Arg Ala			
1665	1670	1675	1680
Leu Cys Leu Val Asp Val Lys Lys Val Lys Gln Ser Leu Ala Gln Ser			
	1685	1690	1695
His Leu Pro Ala Gln Pro Asp Ile Ser Pro Asn Ile Phe Glu Ala Val			
	1700	1705	1710
Lys Gly Cys His Leu Phe Gly Ala Gly Lys Ile Glu Asn Gly Leu Cys			
	1715	1720	1725
Ile Cys Ala Ala Met Pro Ser Lys Val Val Ile Leu Arg Tyr Asn Glu			
	1730	1735	1740
Asn Leu Ser Lys Tyr Cys Ile Arg Lys Glu Ile Glu Thr Ser Glu Pro			
1745	1750	1755	1760
Cys Ser Cys Ile His Phe Thr Asn Tyr Ser Ile Leu Ile Gly Thr Asn			
	1765	1770	1775

- 10/84 -

Fig. 2 (continued)

Lys Phe Tyr Glu Ile Asp Met Lys Gln Tyr Thr Leu Glu Glu Phe Leu	1780	1785	1790
Asp Lys Asn Asp His Ser Leu Ala Pro Ala Val Phe Ala Ala Ser Ser	1795	1800	1805
Asn Ser Phe Pro Val Ser Ile Val Gln Val Asn Ser Ala Gly Gln Arg	1810	1815	1820
Glu Glu Tyr Leu Leu Cys Phe His Glu Phe Gly Val Phe Val Asp Ser	1825	1830	1835
Tyr Gly Arg Arg Ser Arg Thr Asp Asp Leu Lys Trp Ser Arg Leu Pro	1845	1850	1855
Leu Ala Phe Ala Tyr Arg Glu Pro Tyr Leu Phe Val Thr His Phe Asn	1860	1865	1870
Ser Leu Glu Val Ile Glu Ile Gln Ala Arg Ser Ser Ala Gly Thr Pro	1875	1880	1885
Ala Arg Ala Tyr Leu Asp Ile Pro Asn Pro Arg Tyr Leu Gly Pro Ala	1890	1895	1900
Ile Ser Ser Gly Ala Ile Tyr Leu Ala Ser Ser Tyr Gln Asp Lys Leu	1905	1910	1915
Arg Val Ile Cys Cys Lys Gly Asn Leu Val Lys Glu Ser Gly Thr Glu	1925	1930	1935
His His Arg Gly Pro Ser Thr Ser Arg Ser Ser Pro Asn Lys Arg Gly	1940	1945	1950
Pro Pro Thr Tyr Asn Glu His Ile Thr Lys Arg Val Ala Ser Ser Pro	1955	1960	1965
Ala Pro Pro Glu Gly Pro Ser His Pro Arg Glu Pro Ser Thr Pro His	1970	1975	1980
Arg Tyr Arg Glu Gly Arg Thr Glu Leu Arg Arg Asp Lys Ser Pro Gly	1985	1990	1995
Arg Pro Leu Glu Arg Glu Lys Ser Pro Gly Arg Met Leu Ser Thr Arg	2005	2010	2015
Arg Glu Arg Ser Pro Gly Arg Leu Phe Glu Asp Ser Ser Arg Gly Arg	2020	2025	2030
Leu Pro Ala Gly Ala Val Arg Thr Pro Leu Ser Gln Val Asn Lys Val	2035	2040	2045
Trp Asp Gln Ser Ser Val	2050		

- 11/84 -

Fig. 3

Met	Leu	Lys	Phe	Lys	Tyr	Gly	Val	Arg	Asn	Pro	Pro	Glu	Ala	Ser	Ala	1	5	10	15
Ser	Glu	Pro	Ile	Ala	Ser	Arg	Ala	Ser	Arg	Leu	Asn	Leu	Phe	Phe	Gln	20	25	30	
Gly	Lys	Pro	Pro	Leu	Met	Thr	Gln	Gln	Gln	Met	Ser	Ala	Leu	Ser	Arg	35	40	45	
Glu	Gly	Met	Leu	Asp	Ala	Leu	Phe	Ala	Leu	Phe	Glu	Glu	Cys	Ser	Gln	50	55	60	
Pro	Ala	Leu	Met	Lys	Met	Lys	His	Val	Ser	Ser	Phe	Val	Gln	Lys	Tyr	65	70	75	80
Ser	Asp	Thr	Ile	Ala	Glu	Leu	Arg	Glu	Leu	Gln	Pro	Ser	Ala	Arg	Asp	85	90	95	
Phe	Glu	Val	Arg	Ser	Leu	Val	Gly	Cys	Gly	His	Phe	Ala	Glu	Val	Gln	100	105	110	
Val	Val	Arg	Glu	Lys	Ala	Thr	Gly	Asp	Val	Tyr	Ala	Met	Lys	Ile	Met	115	120	125	
Lys	Lys	Lys	Ala	Leu	Leu	Ala	Gln	Glu	Gln	Val	Ser	Phe	Phe	Glu	Glu	130	135	140	
Glu	Arg	Asn	Ile	Leu	Ser	Arg	Ser	Thr	Ser	Pro	Trp	Ile	Pro	Gln	Leu	145	150	155	160
Gln	Tyr	Ala	Phe	Gln	Asp	Lys	Asn	Asn	Leu	Tyr	Leu	Val	Met	Glu	Tyr	165	170	175	
Gln	Pro	Gly	Gly	Asp	Phe	Leu	Ser	Leu	Leu	Asn	Arg	Tyr	Glu	Asp	Gln	180	185	190	
Leu	Asp	Glu	Ser	Met	Ile	Gln	Phe	Tyr	Leu	Ala	Glu	Leu	Ile	Leu	Ala	195	200	205	
Val	His	Ser	Val	His	Gln	Met	Gly	Tyr	Val	His	Arg	Asp	Ile	Lys	Pro	210	215	220	
Glu	Asn	Ile	Leu	Ile	Asp	Arg	Thr	Gly	Glu	Ile	Lys	Leu	Val	Asp	Phe	225	230	235	240
Gly	Ser	Ala	Ala	Lys	Met	Asn	Ser	Asn	Lys	Val	Asp	Ala	Lys	Leu	Pro	245	250	255	
Ile	Gly	Thr	Pro	Asp	Tyr	Met	Ala	Pro	Glu	Val	Leu	Thr	Val	Met	Asn	260	265	270	
Glu	Asp	Arg	Arg	Gly	Thr	Tyr	Gly	Leu	Asp	Cys	Asp	Trp	Trp	Ser	Val	275	280	285	
Gly	Val	Val	Ala	Tyr	Glu	Met	Val	Tyr	Gly	Lys	Thr	Pro	Phe	Thr	Glu	290	295	300	

- 12/84 -

Fig. 3 (continued)

Gly Thr Ser Ala Arg Thr Phe Asn Asn Ile Met Asn Phe Gln Arg Phe			
305	310	315	320
Leu Lys Phe Pro Asp Asp Pro Lys Val Ser Ser Glu Leu Leu Asp Leu			
	325	330	335
Leu Gln Ser Leu Leu Cys Val Gln Lys Glu Arg Leu Lys Phe Glu Gly			
	340	345	350
Leu Cys Cys His Pro Phe Phe Ala Arg Thr Asp Trp Asn Asn Ile Arg			
	355	360	365
Asn Ser Pro Pro Pro Phe Val Pro Thr Leu Lys Ser Asp Asp Asp Thr			
	370	375	380
Ser Asn Phe Asp Glu Pro Glu Lys Asn Ser Trp Ala Phe Ile Leu Cys			
385	390	395	400
Val Pro Ala Glu Pro Leu Ala Phe Ser Gly Glu Glu Leu Pro Phe Val			
	405	410	415
Gly Phe Ser Tyr Ser Lys Ala Leu Gly Tyr Leu Gly Arg Ser Glu Ser			
	420	425	430
Val Val Ser Ser Leu Asp Ser Pro Ala Lys Val Ser Ser Met Glu Lys			
	435	440	445
Lys Leu Leu Ile Lys Ser Lys Glu Leu Gln Asp Ser Gln Asp Lys Cys			
	450	455	460
His Lys Met Glu Gln Glu Met Thr Arg Leu His Arg Arg Val Ser Glu			
465	470	475	480
Val Glu Ala Val Leu Ser Gln Lys Glu Val Glu Leu Lys Ala Ser Glu			
	485	490	495
Thr Gln Arg Ser Leu Leu Glu Gln Asp Leu Ala Thr Tyr Ile Thr Glu			
	500	505	510
Cys Ser Ser Leu Lys Arg Ser Leu Glu Gln Ala Arg Met Glu Val Ser			
	515	520	525
Gln Glu Asp Asp Lys Ala Leu Gln Leu Leu His Asp Ile Arg Glu Gln			
	530	535	540
Ser Arg Lys Leu Gln Glu Ile Lys Glu Gln Glu Tyr Gln Ala Gln Val			
545	550	555	560
Glu Glu Met Arg Leu Met Met Asn Gln Leu Glu Glu Asp Leu Val Ser			
	565	570	575
Ala Arg Arg Arg Ser Asp Leu Tyr Glu Ser Glu Leu Arg Glu Ser Arg			
	580	585	590

- 13/84 -

Fig. 3 (continued)

Leu	Ala	Ala	Glu	Glu	Phe	Lys	Arg	Lys	Ala	Asn	Glu	Cys	Gln	His	Lys	595	600	605
Leu	Met	Lys	Ala	Lys	Asp	Gln	Gly	Lys	Pro	Glu	Val	Gly	Glu	Tyr	Ser	610	615	620
Lys	Leu	Glu	Lys	Ile	Asn	Ala	Glu	Gln	Gln	Leu	Lys	Ile	Gln	Glu	Leu	625	630	635
Gln	Glu	Lys	Leu	Glu	Lys	Ala	Val	Lys	Ala	Ser	Thr	Glu	Ala	Thr	Glu	645	650	655
Leu	Leu	Gln	Asn	Ile	Arg	Gln	Ala	Lys	Glu	Arg	Ala	Glu	Arg	Glu	Leu	660	665	670
Glu	Lys	Leu	His	Asn	Arg	Glu	Asp	Ser	Ser	Glu	Gly	Ile	Lys	Lys	Lys	675	680	685
Leu	Val	Glu	Ala	Glu	Glu	Arg	Arg	His	Ser	Leu	Glu	Asn	Lys	Val	Lys	690	695	700
Arg	Leu	Glu	Thr	Met	Glu	Arg	Arg	Glu	Asn	Arg	Leu	Lys	Asp	Asp	Ile	705	710	715
Gln	Thr	Lys	Ser	Glu	Gln	Ile	Gln	Gln	Met	Ala	Asp	Lys	Ile	Leu	Glu	725	730	735
Leu	Glu	Glu	Lys	His	Arg	Glu	Ala	Gln	Val	Ser	Ala	Gln	His	Leu	Glu	740	745	750
Val	His	Leu	Lys	Gln	Lys	Glu	Gln	His	Tyr	Glu	Glu	Lys	Ile	Lys	Val	755	760	765
Leu	Asp	Asn	Gln	Ile	Lys	Lys	Asp	Leu	Ala	Asp	Lys	Glu	Ser	Leu	Glu	770	775	780
Asn	Met	Met	Gln	Arg	His	Glu	Glu	Glu	Ala	His	Glu	Lys	Gly	Lys	Ile	785	790	795
Leu	Ser	Glu	Gln	Lys	Ala	Met	Ile	Asn	Ala	Met	Asp	Ser	Lys	Ile	Arg	805	810	815
Ser	Leu	Glu	Gln	Arg	Ile	Val	Glu	Leu	Ser	Glu	Ala	Asn	Lys	Leu	Ala	820	825	830
Ala	Asn	Ser	Ser	Leu	Phe	Thr	Gln	Arg	Asn	Met	Lys	Ala	Gln	Glu	Glu	835	840	845
Met	Ile	Ser	Glu	Leu	Arg	Gln	Gln	Lys	Phe	Tyr	Leu	Glu	Thr	Gln	Ala	850	855	860
Gly	Lys	Leu	Glu	Ala	Gln	Asn	Arg	Lys	Leu	Glu	Glu	Gln	Leu	Glu	Lys	865	870	875
Ile	Ser	His	Gln	Asp	His	Ser	Asp	Lys	Ser	Arg	Leu	Leu	Glu	Leu	Glu	885	890	895

- 14/84 -

Fig. 3 (continued)

Thr Arg Leu Arg Glu Val Ser Leu Glu His Glu Glu Gln Lys Leu Glu			
900	905	910	
Leu Lys Arg Gln Leu Thr Glu Leu Gln Leu Ser Leu Gln Glu Arg Glu			
915	920	925	
Ser Gln Leu Thr Ala Leu Gln Ala Ala Arg Ala Ala Leu Glu Ser Gln			
930	935	940	
Leu Arg Gln Ala Lys Thr Glu Leu Glu Glu Thr Thr Ala Glu Ala Glu			
945	950	955	960
Glu Glu Ile Gln Ala Leu Thr Ala His Arg Asp Glu Ile Gln Arg Lys			
965	970	975	
Phe Asp Ala Leu Arg Asn Ser Cys Thr Val Ile Thr Asp Leu Glu Glu			
980	985	990	
Gln Leu Asn Gln Leu Thr Glu Asp Asn Ala Glu Leu Asn Asn Gln Asn			
995	1000	1005	
Phe Tyr Leu Ser Lys Gln Leu Asp Glu Ala Ser Gly Ala Asn Asp Glu			
1010	1015	1020	
Ile Val Gln Leu Arg Ser Glu Val Asp His Leu Arg Arg Glu Ile Thr			
1025	1030	1035	1040
Glu Arg Glu Met Gln Leu Thr Ser Gln Lys Gln Thr Met Glu Ala Leu			
1045	1050	1055	
Lys Thr Thr Cys Thr Met Leu Glu Glu Gln Val Leu Asp Leu Glu Ala			
1060	1065	1070	
Leu Asn Asp Glu Leu Leu Glu Lys Glu Arg Gln Trp Glu Ala Trp Arg			
1075	1080	1085	
Ser Val Leu Gly Asp Glu Lys Ser Gln Phe Glu Cys Arg Val Arg Glu			
1090	1095	1100	
Leu Gln Arg Met Leu Asp Thr Glu Lys Gln Ser Arg Ala Arg Ala Asp			
1105	1110	1115	1120
Gln Arg Ile Thr Glu Ser Arg Gln Val Val Glu Leu Ala Val Lys Glu			
1125	1130	1135	
His Lys Ala Glu Ile Leu Ala Leu Gln Gln Ala Leu Lys Glu Gln Lys			
1140	1145	1150	
Leu Lys Ala Glu Ser Leu Ser Asp Lys Leu Asn Asp Leu Glu Lys Lys			
1155	1160	1165	
His Ala Met Leu Glu Met Asn Ala Arg Ser Leu Gln Gln Lys Leu Glu			
1170	1175	1180	

- 15/84 -

Fig. 3 (continued)

Thr Glu Arg Glu Leu Lys Gln Arg Leu Leu Glu Glu Gln Ala Lys Leu			
1185	1190	1195	1200
Gln Gln Gln Met Asp Leu Gln Lys Asn His Ile Phe Arg Leu Thr Gln			
	1205	1210	1215
Gly Leu Gln Glu Ala Leu Asp Arg Ala Asp Leu Leu Lys Thr Glu Arg			
	1220	1225	1230
Ser Asp Leu Glu Tyr Gln Leu Glu Asn Ile Gln Val Leu Tyr Ser His			
	1235	1240	1245
Glu Lys Val Lys Met Glu Gly Thr Ile Ser Gln Gln Thr Lys Leu Ile			
	1250	1255	1260
Asp Phe Leu Gln Ala Lys Met Asp Gln Pro Ala Lys Lys Lys Lys Val			
1265	1270	1275	1280
Pro Leu Gln Tyr Asn Glu Leu Lys Leu Ala Leu Glu Lys Glu Lys Ala			
	1285	1290	1295
Arg Cys Ala Glu Leu Glu Glu Ala Leu Gln Lys Thr Arg Ile Glu Leu			
	1300	1305	1310
Arg Ser Ala Arg Glu Glu Ala Ala His Arg Lys Ala Thr Asp His Pro			
	1315	1320	1325
His Pro Ser Thr Pro Ala Thr Ala Arg Gln Gln Ile Ala Met Ser Ala			
	1330	1335	1340
Ile Val Arg Ser Pro Glu His Gln Pro Ser Ala Met Ser Leu Leu Ala			
1345	1350	1355	1360
Pro Pro Ser Ser Arg Arg Lys Glu Ser Ser Thr Pro Glu Glu Phe Ser			
	1365	1370	1375
Arg Arg Leu Lys Glu Arg Met His His Asn Ile Pro His Arg Phe Asn			
	1380	1385	1390
Val Gly Leu Asn Met Arg Ala Thr Lys Cys Ala Val Cys Leu Asp Thr			
	1395	1400	1405
Val His Phe Gly Arg Gln Ala Ser Lys Cys Leu Glu Cys Gln Val Met			
	1410	1415	1420
Cys His Pro Lys Cys Ser Thr Cys Leu Pro Ala Thr Cys Gly Leu Pro			
1425	1430	1435	1440
Ala Glu Tyr Ala Thr His Phe Thr Glu Ala Phe Cys Arg Asp Lys Met			
	1445	1450	1455
Asn Ser Pro Gly Leu Gln Ser Lys Glu Pro Gly Ser Ser Leu His Leu			
	1460	1465	1470

- 16/84 -

Fig. 3 (continued)

Glu Gly Trp Met Lys Val Pro Arg Asn Asn Lys Arg Gly Gln Gln Gly	1475	1480	1485
Trp Asp Arg Lys Tyr Ile Val Leu Glu Gly Ser Lys Val Leu Ile Tyr	1490	1495	1500
Asp Asn Glu Ala Arg Glu Ala Gly Gln Arg Pro Val Glu Glu Phe Glu	1505	1510	1515
Leu Cys Leu Pro Asp Gly Asp Val Ser Ile His Gly Ala Val Gly Ala	1525	1530	1535
Ser Glu Leu Ala Asn Thr Ala Lys Ala Asp Val Pro Tyr Ile Leu Lys	1540	1545	1550
Met Glu Ser His Pro His Thr Thr Cys Trp Pro Gly Arg Thr Leu Tyr	1555	1560	1565
Leu Leu Ala Pro Ser Phe Pro Asp Lys Gln Arg Trp Val Thr Ala Leu	1570	1575	1580
Glu Ser Val Val Ala Gly Gly Arg Val Ser Arg Glu Lys Ala Glu Ala	1585	1590	1595
Asp Ala Lys Leu Leu Gly Asn Ser Leu Leu Lys Leu Glu Gly Asp Asp	1605	1610	1615
Arg Leu Asp Met Asn Cys Thr Leu Pro Phe Ser Asp Gln Val Val Leu	1620	1625	1630
Val Gly Thr Glu Glu Gly Leu Tyr Ala Leu Asn Val Leu Lys Asn Ser	1635	1640	1645
Leu Thr His Ile Pro Gly Ile Gly Ala Val Phe Gln Ile Tyr Ile Ile	1650	1655	1660
Lys Asp Leu Glu Lys Leu Leu Met Ile Ala Gly Glu Glu Arg Ala Leu	1665	1670	1675
Cys Leu Val Asp Val Lys Lys Val Lys Gln Ser Leu Ala Gln Ser His	1685	1690	1695
Leu Pro Ala Gln Pro Asp Val Ser Pro Asn Ile Phe Glu Ala Val Lys	1700	1705	1710
Gly Cys His Leu Phe Ala Ala Gly Lys Ile Glu Asn Ser Leu Cys Ile	1715	1720	1725
Cys Ala Ala Met Pro Ser Lys Val Val Ile Leu Arg Tyr Asn Asp Asn	1730	1735	1740
Leu Ser Lys Tyr Cys Ile Arg Lys Glu Ile Glu Thr Ser Glu Pro Cys	1745	1750	1755
			1760

- 17/84 -

Fig. 3 (continued)

Ser Cys Ile His Phe Thr Asn Tyr Ser Ile Leu Ile Gly Thr Asn Lys			
1765	1770	1775	
Phe Tyr Glu Ile Asp Met Lys Gln Tyr Thr Leu Asp Glu Phe Leu Asp			
1780	1785	1790	
Lys Asn Asp His Ser Leu Ala Pro Ala Val Phe Ala Ser Ser Ser Asn			
1795	1800	1805	
Ser Phe Pro Val Ser Ile Val Gln Ala Asn Ser Ala Gly Gln Arg Glu			
1810	1815	1820	
Glu Tyr Leu Leu Cys Phe His Glu Phe Gly Val Phe Val Asp Ser Tyr			
1825	1830	1835	1840
Gly Arg Arg Ser Arg Thr Asp Asp Leu Lys Trp Ser Arg Leu Pro Leu			
1845	1850	1855	
Ala Phe Ala Tyr Arg Glu Pro Tyr Leu Phe Val Thr His Phe Asn Ser			
1860	1865	1870	
Leu Glu Val Ile Glu Ile Gln Ala Arg Ser Ser Leu Gly Ser Pro Ala			
1875	1880	1885	
Arg Ala Tyr Leu Glu Ile Pro Asn Pro Arg Tyr Leu Gly Pro Ala Ile			
1890	1895	1900	
Ser Ser Gly Ala Ile Tyr Leu Ala Ser Ser Tyr Gln Asp Lys Leu Arg			
1905	1910	1915	1920
Val Ile Cys Cys Lys Gly Asn Leu Val Lys Glu Ser Gly Thr Glu Gln			
1925	1930	1935	
His Arg Val Pro Ser Thr Ser Arg Ser Ser Pro Asn Lys Arg Gly Pro			
1940	1945	1950	
Pro Thr Tyr Asn Glu His Ile Thr Lys Arg Val Ala Ser Ser Pro Ala			
1955	1960	1965	
Pro Pro Glu Gly Pro Ser His Pro Arg Glu Pro Ser Thr Pro His Arg			
1970	1975	1980	
Tyr Arg Asp Arg Glu Gly Arg Thr Glu Leu Arg Arg Asp Lys Ser Pro			
1985	1990	1995	2000
Gly Arg Pro Leu Glu Arg Glu Lys Ser Pro Gly Arg Met Leu Ser Thr			
2005	2010	2015	
Arg Arg Glu Arg Ser Pro Gly Arg Leu Phe Glu Asp Ser Ser Arg Gly			
2020	2025	2030	
Arg Leu Pro Ala Gly Ala Val Arg Thr Pro Leu Ser Gln Val Asn Lys			
2035	2040	2045	
Val Trp Asp Gln Ser Ser Val			
2050	2055		

- 18/84 -

Fig. 4

atgttgaagt	tcaaatatgg	agcgcggaat	cctttggatg	ctggtgctgc	tgaacccatt	60
gccagccggg	cctccaggct	gaatctgttc	ttccagggga	aaccaccctt	tatgactcaa	120
cagcagatgt	ctcctctttc	ccgagaaggg	atattagatg	ccctctttgt	tctctttgaa	180
gaatgcagtc	agcctgctct	gatgaagatt	aagcacgtga	gcaactttgt	ccggaagtat	240
tccgacacca	tagctgagtt	acaggagctc	cagccttcgg	caaaggactt	cgaagtcaga	300
agtctttag	gttgtgggtca	ctttgctgaa	gtgcaggtgg	taagagagaa	agcaaccggg	360
gacatctatg	ctatgaaagt	gatgaagaag	aaggctttat	tggcccagga	gcaggtttca	420
ttttttgagg	aagagcggaa	catattatct	cgaagcacia	gcccgtggat	cccccaatta	480
cagtatgcct	ttcaggacaa	aaatcacctt	tatctggtca	tggaatatca	gcctggaggg	540
gacttgctgt	cacttttgaa	tagatatgag	gaccagttag	atgaaaacct	gatacagttt	600
tacctagctg	agctgatttt	ggctgttcac	agcgttcac	tgatgggata	cgtgcatcga	660
gacatcaagc	ctgagaacat	tctcgttgac	cgcacaggac	acatcaagct	ggtggatttt	720
ggatctgccg	cgaaaatgaa	ttcaaacaag	atggtgaatg	ccaaactccc	gattgggacc	780
ccagattaca	tggctcctga	agtgtgact	gtgatgaacg	gggatggaaa	aggcacctac	840
ggcctggact	gtgactgggtg	gtcagtgggc	gtgattgcct	atgagatgat	ttatgggaga	900
tcccccttcg	cagaggggaa	ctctgccaga	accttcaata	acattatgaa	ttccagcgg	960
tttttgaaat	ttccagatga	ccccaaagtg	agcagtgact	ttcttgatct	gattcaaagc	1020
ttgttgctgcg	gccagaaaga	gagactgaag	tttgaaggtc	tttgctgccca	tcctttcttc	1080
tctaaaattg	actggaacaa	cattcgtaac	tctcctcccc	ccttcggttc	cacctcaag	1140
tctgacgatg	acacctccaa	ttttgatgaa	ccagagaaga	attcgtgggt	ttcatcctct	1200
ccgtgccagc	tgagcccctc	aggcttctcg	ggtgaagaac	tgccgtttgt	ggggttttcg	1260
tacagcaagg	cactggggat	tcttggtaga	tctgagtctg	ttgtgtcggg	tctggactcc	1320
cctgccaaaga	ctagctccat	ggaaaagaaa	cttctcatca	aaagcaaaga	gctacaagac	1380
tctcaggaca	agtgtcacaa	gatggagcag	gaaatgaccc	ggttacatcg	gagagtgtca	1440
gaggtggagg	ctgtgcttag	tcagaaggag	gtggagctga	aggcctctga	gactcagaga	1500
tccctcctgg	agcaggacct	tgctacctac	atcacagaat	gcagtagctt	aaagcgaagt	1560
ttggagcaag	cacggatgga	ggtgtcccag	gaggatgaca	aagcactgca	gcttctccat	1620
gatatcagag	agcagagccg	gaagctccaa	gaaatcaaag	agcaggagta	ccaggctcaa	1680
gtggaagaaa	tgaggttgat	gatgaatcag	ttggaagagg	atcttgtctc	agcaagaaga	1740
cggagtgatc	tctacgaatc	tgagctgaga	gagtctcggc	ttgctgctga	agaattcaag	1800
cggaaagcga	cagaatgtca	gcataaactg	ttgaaggcta	aggatcaagg	gaagcctgaa	1860
gtgggagaat	atgcgaaact	ggagaagatc	aatgctgagc	agcagctcaa	aattcaggag	1920
ctccaagaga	aactggagaa	ggctgtaaaa	gccagcacgg	aggccaccga	gctgctgcag	1980
aatatccgcc	aggcaaagga	gcgagccgag	aggagctgg	agaagctgca	gaaccgagag	2040
gattcttctg	aaggcatcag	aaagaagctg	gtggaagctg	aggaacgccg	ccattctctg	2100
gagaacaagg	taaagagact	agagaccatg	gagcgtagag	aaaacagact	gaaggatgac	2160
atccagacaa	aatcccaaca	gatccagcag	atggctgata	aaattctgga	gctcgaagag	2220
aaacatcggg	aggcccaagt	ctcagcccag	cacctagaag	tgcacctgaa	acagaaagag	2280

- 19/84 -

Fig. 4 (continued)

cagcactatg	aggaaaagat	taaagtgttg	gacaatcaga	taaagaaaga	cctggctgac	2340
aaggagacac	tggagaacat	gatgcagaga	cacgaggagg	aggcccatga	gaagggcaaa	2400
attctcagcg	aacagaaggc	gatgatcaat	gctatggatt	ccaagatcag	atccctggaa	2460
cagaggattg	tggaaactgtc	tgaagccaat	aaacttgcag	caaatagcag	tctttttacc	2520
caaaggaaca	tgaaggccca	agaagagatg	atctctgaac	tcaggcaaca	gaaattttac	2580
ctggagacac	aggctgggaa	gttggaggcc	cagaaccgaa	aactggagga	gcagctggag	2640
aagatcagcc	accaagacca	cagtgcacaag	aatcggctgc	tggaaactgga	gacaagattg	2700
cgggaggtca	gtctagagca	cgaggagcag	aaactggagc	tcaagcgcca	gctcacagag	2760
ctacagctct	ccctgcagga	gcgcgagtc	cagttgacag	ccctgcaggc	tgcacgggcg	2820
gccctggaga	gccagcttcg	ccaggcgaag	acagagctgg	aagagaccac	agcagaagct	2880
gaagaggaga	tccaggcact	cacggcacat	agagatgaaa	tccagcgcaa	atctgatgct	2940
cttcgtaaca	gctgtactgt	aatcacagac	ctggaggagc	agctaaacca	gctgaccgag	3000
gacaacgctg	aactcaacaa	ccaaaacttc	tacttgtcca	aacaactcga	tgaggcttct	3060
ggcgccaacg	acgagattgt	acaactgcga	agtgaagtgg	accatctccg	ccgggagatc	3120
acggaacgag	agatgcagct	taccagccag	aagcaaacga	tggaggctct	gaagaccacg	3180
tgcaccatgc	tggaggaaca	ggcatggtat	ttggaggccc	taaacgatga	gctgctagaa	3240
aaagagcggc	agtgggaggc	ctggaggagc	gtcctgggtg	atgagaaatc	ccagtttgag	3300
tgtcgggttc	gagagctgca	gagaatgctg	gacaccgaga	aacagagcag	ggcgagagcc	3360
gatcagcggg	tcaccgagtc	tcgccaggtg	gtggagctgg	cagtgaagga	gcacaaggct	3420
gagattctcg	ctctgcagca	ggctctcaaa	gagcagaagc	tgaaggccga	gagcctctct	3480
gacaagctca	atgacctgga	gaagaagcat	gctatgcttg	aaatgaatgc	ccgaagctta	3540
cagcagaagc	tggagactga	acgagagctc	aaacagaggc	ttctggaaga	gcaagccaaa	3600
ttacagcagc	agatggacct	gcagaaaaat	cacattttcc	gtctgactca	aggactgcaa	3660
gaagctctag	atcgggctga	tctactgaag	acagaaagaa	gtgacttgga	gtatcagctg	3720
gaaaacattc	aggttctcta	ttctcatgaa	aagggtgaaa	tggaaaggac	tatttctcaa	3780
caaaccaaaac	tcattgattt	tctgcaagcc	aaaatggacc	aacctgctaa	aaagaaaaag	3840
gttcctctgc	agtacaatga	gctgaagctg	gccctggaga	aggagaaagc	tcgctgtgca	3900
gagctagagg	aagcccttca	gaagaccgcg	atcgagctcc	ggcccgcccg	ggaggaagct	3960
gccaccgca	aagcaacgga	ccaccacac	ccatccacgc	cagccaccgc	gaggcagcag	4020
atcgccatgt	ccgccatcgt	gcggtcgcca	gagcaccagc	ccagtgccat	gagcctgctg	4080
gccccgccat	ccagccgcag	aaaggagtct	tcaactccag	aggaatttag	tcggcgctct	4140
aaggaacgca	tgcaccacaa	tattcctcac	cgattcaacg	taggactgaa	catgcgagcc	4200
acaaagtgtg	ctgtgtgtct	ggataccgtg	cactttggac	gccaggcatc	caaagtgtct	4260
gaatgtcagg	tgatgtgtca	ccccaaagtgc	tccacgtgct	tgccagccac	ctgcggcttg	4320
cctgctgaat	atgccacaca	cttcaccgag	gccttctgcc	gtgacaaaat	gaactcccca	4380
ggtctccaga	ccaaggagcc	cagcagcagc	ttgcacctgg	aagggtggat	gaaggtgccc	4440
aggaataaca	aacgaggaca	gcaaggctgg	gacaggaagt	acattgtcct	ggagggatca	4500
aaagtcctca	tttatgacaa	tgaagccaga	gaagctggac	agaggccggt	ggaagaatct	4560

- 20/84 -

Fig. 4 (continued)

gagctgtgcc	ttcccgaagg	ggatgtatct	attcatgggtg	ccgttggtgc	ttccgaactc	4620
gcaaatacag	ccaaagcaga	tgtcccatac	atactgaaga	tggaatctca	cccgcacacc	4680
acctgctggc	ccgggagaac	cctctacttg	ctagctccca	gcttccttga	caaacagcgc	4740
tgggtcaccg	ccttagaatc	agttgtcgca	gggtgggagag	tttctaggga	aaaagcagaa	4800
gctgatgcta	aactgcttgg	aaactccctg	ctgaaactgg	aaggtgatga	ccgtctagac	4860
atgaactgca	cgctgccctt	cagtgaccag	gtggtgttgg	tgggcaccga	ggaagggctc	4920
tacgccctga	atgtcttgaa	aaactcccta	acccatgtcc	caggaattgg	agcagtcttc	4980
caaattttata	ttatcaagga	cctggagaag	ctactcatga	tagcaggaga	agagcgggca	5040
ctgtgtcttg	tggacgtgaa	gaaagtgaaa	cagtccttgg	cccagtccca	cctgcctgcc	5100
cagccccgaca	tctcaccocaa	cattttttgaa	gctgtcaagg	gctgccactt	gtttggggca	5160
ggcaagattg	agaacgggct	ctgcatctgt	gcagccatgc	ccagcaaagt	cgtcattctc	5220
cgctacaacg	aaaacctcag	caaatactgc	atccggaaag	agatagagac	ctcagagccc	5280
tgcagctgta	tccacttcac	caattacagt	atcctcattg	gaaccaataa	attctacgaa	5340
atcgacatga	agcagtacac	gctcgaggaa	ttcctggata	agaatgacca	ttccttggca	5400
cctgctgtgt	ttgccgcctc	ttccaacagc	ttcctgtct	caatcgtgca	gggtgaacagc	5460
gcagggcgagc	gagaggagta	cttgctgtgt	ttccacgaat	ttggagtgtt	cgtggattct	5520
tacggaagac	gtagccgcac	agacgatctc	aagtggagtc	gcttaccttt	ggcctttgcc	5580
tacagagaac	cctatctgtt	tgtgaccac	ttcaactcac	tcgaagtaat	tgagatccag	5640
gcacgctcct	cagcaggagc	ccctgcccga	gcgtacctgg	acatcccga	cccgcgctac	5700
ctgggcccctg	ccatttctct	aggagcgatt	tacttggcgt	cctcatacca	ggataaatta	5760
agggtcattt	gctgcaaggg	aaacctcgtg	aaggagtccg	gcactgaaca	ccaccggggc	5820
ccgtccacct	cccgcagcag	ccccaaacaag	cgaggccac	ccacgtacaa	cgagcacatc	5880
accaagcgcg	tggcctccag	cccagcgccg	cccgaaggcc	ccagccaccc	gcgagagcca	5940
agcacacccc	accgctaccg	cgagggggcg	accgagctgc	gcagggacaa	gtctcctggc	6000
cgccccctgg	agcgagagaa	gtcccccgcc	cggatgctca	gcacgcggag	agagcggctc	6060
cccgggaggc	tgtttgaaga	cagcagcagg	ggccggctgc	ctgcgggagc	cgtgaggacc	6120
ccgctgtccc	aggtgaacaa	ggtctgggac	cagtcttcag	tataaatctc	agccagaaaa	6180
accaactcct	catcttgatc	tgcaggaaaa	caccaaacac	actatggaac	tctgctgatg	6240
gggacccaag	cgcccacgtg	ctcagccacc	ctctggctca	gcggggccca	gacccacctc	6300
ggcacggaca	cccctgtctc	caggaggggc	aggtggctga	ggctcttogg	agctgtcagc	6360
gcccgggtgcc	tgccttgggc	acctccctgc	agtcattctc	ttgcactttg	ttactctttc	6420
aaagcattca	caaacttttg	tacctagctc	tagcctgtac	cagttagtct	atcaaaggaa	6480
accaaccggg	atgctaacaa	caacatgggt	agaatcctaa	ttagctactt	taagatccta	6540
ggattggttg	gtttttcttt	ttttttctc	tttgtttctt	tccttttttt	tttttttttt	6600
taagacaaca	gaattottaa	tagatttgaa	tagcgacgta	ttcctgttg	tagtcatttt	6660
tagctcgacc	acatcatcag	gtctttgcc	ccgaggcata	gtgtagaaca	gtcccgttca	6720
gttggccaac	ctcccgcagc	caagtaggtt	catccttgtt	cctgttcatt	ctcatagatg	6780

- 21/84 -

Fig. 4 (continued)

gcccctgcttt	ccccaggggtg	acatcgtagc	caaatgttta	ctgttttcat	tgccttttat	6840
ggccttgacg	acttcccctc	ccaccagctg	agaatgtatg	gaggtcatcg	gggcctcagc	6900
tgggaggcag	tgacttgggg	ccaagggacc	tcgagacgct	ttccttcccc	accccccagc	6960
gtcatctccc	cagcctgctg	ttcccgcctt	ccatatagct	ttggccagga	aagcatgcaa	7020
tagacttgct	cggagcccag	cactcctggg	tctcggggtc	ggggagggga	cgggggaccc	7080
cacttccttg	tctgtgacgg	cgtgttgctc	cccactctgg	gatggggaag	aggcccgtcg	7140
ggagttctgc	atggcagttc	actgcatgtg	ctgccccctt	gggttgctct	gccaatgtat	7200
taataccatc	ccatagctcc	tgccaaatcg	agaccctctg	acgacttgcc	gactaactgg	7260
ccaccacaag	ctgcagtctg	tagcactgaa	caaacaaaaa	acaaaacgct	caagccttac	7320
gaccagagaa	ggatttcagc	aaaccaccac	ctcccactca	gtgtcccctc	caaacttcac	7380
acttcccctg	ctgcagagga	tgactctggt	cacacccaat	ccagcgcggg	tctaccccac	7440
gaaactgtga	ctttccaaat	gagcctttcc	ctagggctag	acctaagacc	aggaagtgtg	7500
agaaagcagc	cgcagctcaa	ctcttccagc	tccgccaggg	ttgggaagtc	cttaggtgca	7560
gtgcggtctc	cactgggtct	gcggaccctc	ctattagagt	acgaaattcc	tggcaactgg	7620
tatagaacca	acctagaggg	tttgacgttg	gcaagctaac	tcgcggcctt	atttctgcct	7680
ttaatctccc	acaaggcatc	tgttgctttg	ggtcctccac	gactcttagg	cccgcctcaa	7740
caaccagggc	acctcctagg	taggctcaaa	ggtagaccgg	tttccaccgc	agcagggtgaa	7800
catgaccgtg	ttttcaactg	tgtccacagt	tcagatccct	ttccagattg	caacctggcc	7860
tgcattcccag	ctccttctct	ctcgtgtctt	aacctaagtg	ctttcttggt	tgaaacgcct	7920
acaaacctcc	atgtggtagc	tcctttggca	aatgtcctgc	tgtggcgttt	tatgtgttgc	7980
ttggagtctg	tggggtcgta	ctccctcccc	tcccgctccc	agggcagatt	tgattgaatg	8040
tttgctgaag	ttttgtctct	tgggtccacag	tatttggaag	ggctactgaa	aatgggtctt	8100
tcagtcttgg	catttcattt	aggatctcca	tgagaaatgg	gcttcttgag	ccctgaaaat	8160
gtatatattg	tgtctcatct	gtgaaactgct	ttctgtata	tagaactagc	tcaaaagact	8220
gtacatattt	acaagaaact	ttatatctgt	aaaaaaaaaa	agaggaaatt	gaattggttt	8280
ctactttttt	attgtaaaag	gtgcattttt	caacacttac	ttttgggttc	aatgggtgta	8340
gttgtggaca	gccatcttca	ctggaggggtg	gggagctccg	tgtgaccacc	aagatgccag	8400
caggatatac	cgtaacacga	aattgctgtc	aaaagcttat	tagcatcaat	caagattcta	8460
ggtctccaaa	agtacaggct	ttttcttcat	tacctttttt	attcagaacg	aggaagagaa	8520
cacaaggaat	gattcaagat	ccaccttgag	aggaatgaac	tttggtgttg	aacaattagt	8580
gaaataaagc	aatgatctaa	act				8603

- 22/84 -

Fig. 5

Val	Leu	Asp	Asn	Gln	Ile	Lys	Lys	Asp	Leu	Ala	Asp	Lys	Glu	Thr	Leu
1				5					10					15	
Glu	Asn	Met	Met	Gln	Arg	His	Glu	Glu	Glu	Ala	His	Glu	Lys	Gly	Lys
			20					25					30		
Ile	Leu	Ser	Glu	Gln	Lys	Ala	Met	Ile	Asn	Ala	Met	Asp	Ser	Lys	Ile
			35				40					45			
Arg	Ser	Leu	Glu	Gln	Arg	Ile	Val	Glu	Leu	Ser	Glu	Ala	Asn	Lys	Leu
			50				55				60				
Ala	Ala	Asn	Ser	Ser	Leu	Phe	Thr	Gln	Arg	Asn	Met	Lys	Ala	Gln	Glu
65					70					75				80	
Glu	Met	Ile	Ser	Glu	Leu	Arg	Gln	Gln	Lys	Phe	Tyr	Leu	Glu	Thr	Gln
				85					90					95	
Ala	Gly	Lys	Leu	Glu	Ala	Gln	Asn	Arg	Lys	Leu	Glu	Glu	Gln	Leu	Glu
			100					105					110		
Lys	Ile	Ser	His	Gln	Asp	His	Ser	Asp	Lys	Asn	Arg	Leu	Leu	Glu	Leu
			115				120					125			
Glu	Thr	Arg	Leu	Arg	Glu	Val	Ser	Leu	Glu	His	Glu	Glu	Gln	Lys	Leu
			130				135					140			
Glu	Leu	Lys	Arg	Gln	Leu	Thr	Glu	Leu	Gln	Leu	Ser	Leu	Gln	Glu	Arg
145					150					155				160	
Glu	Ser	Gln	Leu	Thr	Ala	Leu	Gln	Ala	Ala	Arg	Ala	Ala	Leu	Glu	Ser
				165					170					175	
Gln	Leu	Arg	Gln	Ala	Lys	Thr	Glu	Leu	Glu	Glu	Thr	Thr	Ala	Glu	Ala
			180					185					190		
Glu	Glu	Glu	Ile	Gln	Ala	Leu	Thr	Ala	His	Arg	Asp	Glu	Ile	Gln	Arg
			195				200					205			
Lys	Phe	Asp	Ala	Leu	Arg	Asn	Ser	Cys	Thr	Val	Ile	Thr	Asp	Leu	Glu
			210				215					220			
Glu	Gln	Leu	Asn	Gln	Leu	Thr	Glu	Asp	Asn	Ala	Glu	Leu	Asn	Asn	Gln
225					230					235				240	
Asn	Phe	Tyr	Leu	Ser	Lys	Gln	Leu	Asp	Glu	Ala	Ser	Gly	Ala	Asn	Asp
				245					250					255	
Glu	Ile	Val	Gln	Leu	Arg	Ser	Glu	Val	Asp	His	Leu	Arg	Arg	Glu	Ile
			260					265					270		
Thr	Glu	Arg	Glu	Met	Gln	Leu	Thr	Ser	Gln	Lys	Gln	Thr	Met	Glu	Ala
			275				280					285			
Leu	Lys	Thr	Thr	Cys	Thr	Met	Leu	Glu	Glu	Gln	Val	Met	Asp	Leu	Glu
			290				295					300			

- 23/84 -

Fig. 5 (continued)

Ala Leu Asn Asp Glu Leu Leu Glu Lys Glu Arg Gln Trp Glu Ala Trp			
305	310	315	320
Arg Ser Val Leu Gly Asp Glu Lys Ser Gln Phe Glu Cys Arg Val Arg			
	325	330	335
Glu Leu Gln Arg Met Leu Asp Thr Glu Lys Gln Ser Arg Ala Arg Ala			
	340	345	350
Asp Gln Arg Ile Thr Glu Ser Arg Gln Val Val Glu Leu Ala Val Lys			
	355	360	365
Glu His Lys Ala Glu Ile Leu Ala Leu Gln Gln Ala Leu Lys Glu Gln			
	370	375	380
Lys Leu Lys Ala Glu Ser Leu Ser Asp Lys Leu Asn Asp Leu Glu Lys			
385	390	395	400
Lys His Ala Met Leu Glu Met Asn Ala Arg Ser Leu Gln Gln Lys Leu			
	405	410	415
Glu Thr Glu Arg Glu Leu Lys Gln Arg Leu Leu Glu Glu Gln Ala Lys			
	420	425	430
Leu Gln Gln Gln Met Asp Leu Gln Lys Asn His Ile Phe Arg Leu Thr			
	435	440	445
Gln Gly Leu Gln Glu Ala Leu Asp Arg Ala Asp Leu Leu Lys Thr Glu			
	450	455	460
Arg Ser Asp Leu Glu Tyr Gln Leu Glu Asn Ile Gln Val Leu Tyr Ser			
465	470	475	480
His Glu Lys Val Lys Met Glu Gly Thr Ile Ser Gln Gln Thr Lys Leu			
	485	490	495
Ile Asp Phe Leu Gln Ala Lys Met Asp Gln Pro Ala Lys Lys Lys Lys			
	500	505	510
Val Pro Leu Gln Tyr Asn Glu Leu Lys Leu Ala Leu Glu Lys Glu Lys			
	515	520	525
Ala Arg Cys Ala Glu Leu Glu Glu Ala Leu Gln Lys Thr Arg Ile Glu			
	530	535	540
Leu Arg Ser Ala Arg Glu Glu Ala Ala His Arg Lys Ala Thr Asp His			
545	550	555	560
Pro His Pro Ser Thr Pro Ala Thr Ala Arg Gln Gln Ile Ala Met Ser			
	565	570	575
Ala Ile Val Arg Ser Pro Glu His Gln Pro Ser Ala Met Ser Leu Leu			
	580	585	590
Ala Pro Pro Ser Ser Arg Arg Lys Glu Ser Ser Thr Pro Glu Glu Phe			
	595	600	605

- 24/84 -

Fig. 5 (continued)

Ser	Arg	Arg	Leu	Lys	Glu	Arg	Met	His	His	Asn	Ile	Pro	His	Arg	Phe	610	615	620
Asn	Val	Gly	Leu	Asn	Met	Arg	Ala	Thr	Lys	Cys	Ala	Val	Cys	Leu	Asp	625	630	635
Thr	Val	His	Phe	Gly	Arg	Gln	Ala	Ser	Lys	Cys	Leu	Glu	Cys	Gln	Val	645	650	655
Met	Cys	His	Pro	Lys	Cys	Ser	Thr	Cys	Leu	Pro	Ala	Thr	Cys	Gly	Leu	660	665	670
Pro	Ala	Glu	Tyr	Ala	Thr	His	Phe	Thr	Glu	Ala	Phe	Cys	Arg	Asp	Lys	675	680	685
Met	Asn	Ser	Pro	Gly	Leu	Gln	Thr	Lys	Glu	Pro	Ser	Ser	Ser	Leu	His	690	695	700
Leu	Glu	Gly	Trp	Met	Lys	Val	Pro	Arg	Asn	Asn	Lys	Arg	Gly	Gln	Gln	705	710	715
Gly	Trp	Asp	Arg	Lys	Tyr	Ile	Val	Leu	Glu	Gly	Ser	Lys	Val	Leu	Ile	725	730	735
Tyr	Asp	Asn	Glu	Ala	Arg	Glu	Ala	Gly	Gln	Arg	Pro	Val	Glu	Glu	Phe	740	745	750
Glu	Leu	Cys	Leu	Pro	Asp	Gly	Asp	Val	Ser	Ile	His	Gly	Ala	Val	Gly	755	760	765
Ala	Ser	Glu	Leu	Ala	Asn	Thr	Ala	Lys	Ala	Asp	Val	Pro	Tyr	Ile	Leu	770	775	780
Lys	Met	Glu	Ser	His	Pro	His	Thr	Thr	Cys	Trp	Pro	Gly	Arg	Thr	Leu	785	790	795
Tyr	Leu	Leu	Ala	Pro	Ser	Phe	Pro	Asp	Lys	Gln	Arg	Trp	Val	Thr	Ala	805	810	815
Leu	Glu	Ser	Val	Val	Ala	Gly	Gly	Arg	Val	Ser	Arg	Glu	Lys	Ala	Glu	820	825	830
Ala	Asp	Ala	Lys	Leu	Leu	Gly	Asn	Ser	Leu	Leu	Lys	Leu	Glu	Gly	Asp	835	840	845
Asp	Arg	Leu	Asp	Met	Asn	Cys	Thr	Leu	Pro	Phe	Ser	Asp	Gln	Val	Val	850	855	860
Leu	Val	Gly	Thr	Glu	Glu	Gly	Leu	Tyr	Ala	Leu	Asn	Val	Leu	Lys	Asn	865	870	875
Ser	Leu	Thr	His	Val	Pro	Gly	Ile	Gly	Ala	Val	Phe	Gln	Ile	Tyr	Ile	885	890	895
Ile	Lys	Asp	Leu	Glu	Lys	Leu	Leu	Met	Ile	Ala	Gly	Glu	Glu	Arg	Ala	900	905	910

Leu	Cys	Leu	Val	Asp	Val	Lys	Lys	Val	Lys	Gln	Ser	Leu	Ala	Gln	Ser
915				920				925							
His	Leu	Pro	Ala	Gln	Pro	Asp	Ile	Ser	Pro	Asn	Ile	Phe	Glu	Ala	Val
930				935				940							
Lys	Gly	Cys	His	Leu	Phe	Gly	Ala	Gly	Lys	Ile	Glu	Asn	Gly	Leu	Cys
945				950				955				960			
Ile	Cys	Ala	Ala	Met	Pro	Ser	Lys	Val	Val	Ile	Leu	Arg	Tyr	Asn	Glu
965				970				975							
Asn	Leu	Ser	Lys	Tyr	Cys	Ile	Arg	Lys	Glu	Ile	Glu	Thr	Ser	Glu	Pro
980				985				990							
Cys	Ser	Cys	Ile	His	Phe	Thr	Asn	Tyr	Ser	Ile	Leu	Ile	Gly	Thr	Asn
995				1000				1005							
Lys	Phe	Tyr	Glu	Ile	Asp	Met	Lys	Gln	Tyr	Thr	Leu	Glu	Glu	Phe	Leu
1010				1015				1020							
Asp	Lys	Asn	Asp	His	Ser	Leu	Ala	Pro	Ala	Val	Phe	Ala	Ala	Ser	Ser
1025				1030				1035				1040			
Asn	Ser	Phe	Pro	Val	Ser	Ile	Val	Gln	Val	Asn	Ser	Ala	Gly	Gln	Arg
1045				1050				1055							
Glu	Glu	Tyr	Leu	Leu	Cys	Phe	His	Glu	Phe	Gly	Val	Phe	Val	Asp	Ser
1060				1065				1070							
Tyr	Gly	Arg	Arg	Ser	Arg	Thr	Asp	Asp	Leu	Lys	Trp	Ser	Arg	Leu	Pro
1075				1080				1085							
Leu	Ala	Phe	Ala	Tyr	Arg	Glu	Pro	Tyr	Leu	Phe	Val	Thr	His	Phe	Asn
1090				1095				1100							
Ser	Leu	Glu	Val	Ile	Glu	Ile	Gln	Ala	Arg	Ser	Ser	Ala	Gly	Thr	Pro
1105				1110				1115				1120			
Ala	Arg	Ala	Tyr	Leu	Asp	Ile	Pro	Asn	Pro	Arg	Tyr	Leu	Gly	Pro	Ala
1125				1130				1135							
Ile	Ser	Ser	Gly	Ala	Ile	Tyr	Leu	Ala	Ser	Ser	Tyr	Gln	Asp	Lys	Leu
1140				1145				1150							
Arg	Val	Ile	Cys	Cys	Lys	Gly	Asn	Leu	Val	Lys	Glu	Ser	Gly	Thr	Glu
1155				1160				1165							
His	His	Arg	Gly	Pro	Ser	Thr	Ser	Arg	Ser	Ser	Pro	Asn	Lys	Arg	Gly
1170				1175				1180							
Pro	Pro	Thr	Tyr	Asn	Glu	His	Ile	Thr	Lys	Arg	Val	Ala	Ser	Ser	Pro
1185				1190				1195				1200			
Ala	Pro	Pro	Glu	Gly	Pro	Ser	His	Pro	Arg	Glu	Pro	Ser	Thr	Pro	His
1205				1210				1215							

1220

1225

1230

1235

1240

1245

1250

1255

1260

1265

1270

1275

1280

1285

- 27/84 -

Fig. 6

cagagcaggg	cgagagccga	tcagcggatc	accgagtctc	gccaggtggg	ggagctggca	60
gtgaaggagc	acaaggctga	gattctcgct	ctgcagcagg	ctctcaaaga	gcagaagctg	120
aaggccgaga	gcctctctga	caagctcaat	gacctggaga	agaagcatgc	tatgcttgaa	180
atgaatgccc	gaagcttaca	gcagaagctg	gagactgaac	gagagctcaa	acagaggctt	240
ctggaagagc	aagccaaatt	acagcagcag	atggacctgc	agaaaaatca	cattttccgt	300
ctgactcaag	gactgcaaga	agctctagat	cgggctgata	tactgaagac	agaaagaagt	360
gacttggaat	atcagctgga	aaacattcag	gttctctatt	ctcatgaaaa	ggtgaaaatg	420
gaaggcacta	tttctcaaca	aaccaaactc	attgattttc	tgcaagccaa	aatggaccaa	480
cctgctaaaa	agaaaaaggt	tcctctgcag	tacaatgagc	tgaagctggc	cctggagaag	540
gagaaagctc	gctgtgcaga	gctagaggaa	gcccttcaga	agacccgcat	cgagctccgg	600
tccgcccggg	aggaagctgc	ccaccgcaaa	gcaacggacc	acccacaccc	atccacgcca	660
gccaccgcga	ggcagcagat	cgccatgtct	gccatcgtgc	ggtcgccaga	gcaccagccc	720
agtgccatga	gcctgctggc	cccgccatcc	agccgcagaa	aggagtcttc	aactccagag	780
gaatttagtc	ggcgtcttaa	ggaacgcatg	caccacaata	ttcctcaccg	attcaacgta	840
ggactgaaca	tgcgagccac	aaagtgtgct	gtgtgtctgg	ataccgtgca	ctttggacgc	900
caggcatcca	aatgtctcga	atgtcaggtg	atgtgtcacc	ccaagtgtct	cacgtgcttg	960
ccagccacct	gcggcttgcc	tgctgaatat	gccacacact	tcaccgaggc	cttctgccgt	1020
gacaaaatga	actccccagg	tctccagacc	aaggagccca	gcagcagctt	gcacctggaa	1080
gggtggatga	aggtgcccag	gaataacaaa	cgaggacagc	aaggctggga	caggaagtac	1140
attgtcctgg	agggatcaaa	agtcctcatt	tatgacaatg	aagccagaga	agctggacag	1200
aggccggtgg	aagaatttga	gctgtgcctt	cccgcgggg	atgtatctat	tcatggtgcc	1260
gttgggtgct	ccgaactcgc	aaatacagcc	aaagcagatg	tcccatacat	actgaagatg	1320
gaatctcacc	cgcacaccac	ctgctggccc	gggagaaccc	tctacttgct	agctcccagc	1380
ttccctgaca	aacagcgctg	ggtcaccgcc	ttagaatcag	ttgtcgaggg	tgggagagtt	1440
tctagggaaa	aagcagaagc	tgatgctaaa	ctgcttgaaa	actccctgct	gaaactggaa	1500
ggtgatgacc	gtctagacat	gaactgcacg	ctgcccttca	gtgaccaggt	ggtgttggtg	1560
ggcaccgagg	aagggctcta	cgccctgaat	gtcttgaaaa	actccctaac	ccatgtccca	1620
ggaattggag	cagtcttcca	aatttatatt	atcaaggacc	tggagaagct	actcatgata	1680
gcaggagaag	agcgggcact	gtgtcttggt	gacgtgaaga	aagtgaacaa	gtccctggcc	1740
cagtcccacc	tgccctgcca	gcccgcacatc	tcaccaacaa	tttttgaaagc	tgtcaagggc	1800
tgccacttgt	ttggggcagg	caagattgag	aacgggctct	gcactctgtgc	agccatgccc	1860
agcaaagtgc	tcattctccg	ctacaacgaa	aacctcagca	aatactgcat	ccggaaagag	1920
atagagacct	cagagccctg	cagctgtatc	cacttcacca	attacagtat	cctcattgga	1980
accaataaat	tctacgaaat	cgacatgaag	cagtacacgc	tcgaggaatt	cctggataag	2040
aatgaccatt	ccttggcacc	tgctgtgttt	gccgcctctt	ccaacagctt	cctgtctca	2100
atcgtgcagg	tgaacagcgc	agggcagcga	gaggagtact	tgctgtgttt	ccacgaattt	2160
ggagtgttcg	tggattctta	cggaagacgt	agccgcacag	acgatctcaa	gtggagtccg	2220
ttacctttgg	cctttgccta	cagagaaccc	tatctgtttg	tgaccactt	caactcactc	2280

- 28/84 -

Fig. 6 (continued)

gaagtaattg	agatccaggc	acgctcctca	gcagggaccc	ctgcccagac	gtacctggac	2340
atcccgaacc	cgcgctacct	gggccctgcc	atttcctcag	gagcgattta	cttggcgctcc	2400
tcataccagg	ataaattaag	ggtcatttgc	tgcaagggaa	acctcgtgaa	ggagtccggc	2460
actgaacacc	accggggccc	gtccacctcc	cgcagcagcc	ccaacaagcg	aggcccaccc	2520
acgtacaacg	agcacatcac	caagcgcgtg	gcctccagcc	cagcgccgcc	cgaaggcccc	2580
agccacccgc	gagagccaag	cacaccccac	cgtaccgcg	aggggaggac	cgagctgcgc	2640
agggacaagt	ctcctggccg	ccccctggag	cgagagaagt	cccccgccg	gatgctcagc	2700
acgcggagag	agcgggtccc	cgggaggctg	tttgaagaca	gcagcagggg	ccggctgcct	2760
gcgggagccg	tgaggacccc	gctgtcccag	gtgaacaagg	tctgggacca	gtcttcagta	2820
taaatctcag	ccagaaaaac	caactcctca	tcttgatctg	caggaaaaca	ccaaacacac	2880
tatggaactc	tgctgatggg	gacccaagcg	cccacgtgct	cagccaccct	ctggctcagc	2940
ggggcccaga	cccacctcgg	cacggacacc	cctgtctcca	ggaggggcag	gtggctgagg	3000
ctcttcggag	ctgtcagcgc	ccggtgcctg	ccctgggcac	ctccctgcag	tcctctcttt	3060
gcactttgtt	actctttcaa	agcattcaca	aacttttgta	cctagctcta	gcctgtacca	3120
gttagttcat	caaaggaaac	caaccgggat	gctaacaaca	acatggttag	aatcctaatt	3180
agctacttta	agatcctagg	attggttggg	ttttcttttt	ttttctctct	tgtttctttc	3240
cttttttttt	ttttttttta	agacaacaga	attcttaata	gatttgaata	gcgacgtatt	3300
tcctgttgta	gtcattttta	gctcgaccac	atcatcaggt	ctttgccacc	gaggcatagt	3360
gtagaacagt	ccggctcagt	tggccaacct	ccgcagcca	agtaggttca	tccttgttcc	3420
tgttcattct	catagatggc	cctgctttcc	ccagggtgac	atcgtagcca	aatgtttact	3480
gttttcattg	ccttttatgg	ccttgacgac	ttccccctcc	accagctgag	aatgtatgga	3540
ggtcacgagg	gcctcagctc	ggaggcagtg	acttggggcc	aaggacacct	gagacgcttt	3600
ccttccccac	cccccagcgt	catctcccca	gcctgctggt	cccgttttcc	atatagcttt	3660
ggccaggaaa	gcatgcaata	gacttgctcg	gagcccagca	ctcctgggtc	tcggggtcgg	3720
ggaggggacg	ggggcaccca	cttccttgct	tgtgacggcg	tgttggtccc	cactctggga	3780
tggggaagag	gcccgtcggg	agttctgcat	ggcagttcac	tgcagtgtgt	gcccccttgg	3840
gttgctctgc	caatgtatta	ataccatccc	atagctcctg	ccaaatcgag	accctctgac	3900
gacttgccga	ctaactggcc	accacaagct	gcagtctgta	gcactgaaca	aacaaaaaac	3960
aaaacgctca	agccttacga	ccagagaagg	atttcagcaa	accaccacct	cccactcagt	4020
gtccccctca	aacttcacac	ttccctgcct	gcagaggatg	actctgttca	cacccaatcc	4080
agcgcgggtc	taccccacga	aactgtgact	ttccaaatga	gcctttccct	agggttagac	4140
ctaagaccag	gaagtttgag	aaagcagccg	cagctcaact	cttcagctc	cgccagggtt	4200
gggaagtcct	taggtgcagt	gcggctccca	ctgggtctgc	ggaccctcct	attagagtac	4260
gaaattcctg	gcaactggta	tagaaccaac	ctagaggctt	tgcagttggc	aagctaactc	4320
gcggccttat	ttctgccttt	aatctcccac	aaggcatctg	ttgctttggg	tcctccacga	4380
ctcttaggcc	cgcctcaaca	accaggcac	ctcctaggta	ggctcaaagg	tagaccggtt	4440
tccaccgcag	caggtgaaca	tgaccgtggt	ttcaactgtg	tccacagttc	agatcccttt	4500
ccagattgca	acctggcctg	catcccagct	ccttcctgct	cgtgtcttaa	cctaagtgtc	4560
ttcttgtttg	aaacgcctac	aaacctccat	gtggtagctc	ctttggcaaa	tgtcctgctg	4620

- 29/84 -

Fig. 6 (continued)

tggcggtttta	tgtgttgctt	ggagtctgtg	gggtcgtact	ccctcccctc	ccgtccccag	4680
ggcagatttg	attgaatgtt	tgctgaagtt	ttgtctcttg	gtccacagta	tttggaagg	4740
tcactgaaaa	tggtcttttc	agtcttggca	tttcatttag	gatctccatg	agaaatgggc	4800
ttcttgagcc	ctgaaaatgt	atattgtgtg	tctcatctgt	gaactgcttt	ctgctatata	4860
gaactagctc	aaaagactgt	acatatttac	aagaaacttt	atattcgtaa	aaaaaaaaag	4920
aggaaattga	attggtttct	acttttttat	tgtaaaaggt	gcatttttca	acacttactt	4980
ttggtttcaa	tggtggtagt	tgtggacagc	catcttcact	ggagggtggg	gagctccgtg	5040
tgaccaccaa	gatgccagca	ggatataccg	taacacgaaa	ttgctgtcaa	aagcttatta	5100
gcacatca	agattctagg	tctccaaaag	tacaggcttt	ttcttcatta	ccttttttat	5160
tcagaacgag	gaagagaaca	caaggaatga	ttcaagatcc	accttgagag	gaatgaactt	5220
tgttggtgaa	caattagtga	aataaagcaa	tgatctaaac	t		5261

- 30/84 -

Fig. 7

Met	Ser	Ala	Glu	Val	Arg	Leu	Arg	Gln	Leu	Gln	Gln	Leu	Val	Leu	Asp
1				5					10					15	
Pro	Gly	Phe	Leu	Gly	Leu	Glu	Pro	Leu	Leu	Asp	Leu	Leu	Leu	Gly	Val
			20					25					30		
His	Gln	Glu	Leu	Gly	Ala	Ser	His	Leu	Ala	Gln	Asp	Lys	Tyr	Val	Ala
		35					40					45			
Asp	Phe	Leu	Gln	Trp	Val	Glu	Pro	Ile	Ala	Ala	Arg	Leu	Lys	Glu	Val
	50					55					60				
Arg	Leu	Gln	Arg	Asp	Asp	Phe	Glu	Ile	Leu	Lys	Val	Ile	Gly	Arg	Gly
65				70					75					80	
Ala	Phe	Ser	Glu	Val	Ala	Val	Val	Lys	Met	Lys	Gln	Thr	Gly	Gln	Val
				85				90						95	
Tyr	Ala	Met	Lys	Ile	Met	Asn	Lys	Trp	Asp	Met	Leu	Lys	Arg	Gly	Glu
		100					105						110		
Val	Ser	Cys	Phe	Arg	Glu	Glu	Arg	Asp	Val	Leu	Val	Lys	Gly	Asp	Arg
		115					120					125			
Arg	Trp	Ile	Thr	Gln	Leu	His	Phe	Ala	Phe	Gln	Asp	Glu	Asn	Tyr	Leu
	130					135					140				
Tyr	Leu	Val	Met	Glu	Tyr	Tyr	Val	Gly	Gly	Asp	Leu	Leu	Thr	Leu	Leu
145				150					155					160	
Ser	Lys	Phe	Gly	Glu	Arg	Ile	Pro	Ala	Glu	Met	Ala	Arg	Phe	Tyr	Leu
			165					170					175		
Ala	Glu	Ile	Val	Met	Ala	Ile	Asp	Ser	Val	His	Arg	Leu	Gly	Tyr	Val
		180					185					190			
His	Arg	Asp	Ile	Lys	Pro	Asp	Asn	Ile	Leu	Leu	Asp	Arg	Cys	Gly	His
		195				200					205				
Ile	Arg	Leu	Ala	Asp	Phe	Gly	Ser	Cys	Leu	Lys	Leu	Gln	Pro	Asp	Gly
	210					215					220				
Met	Val	Arg	Ser	Leu	Val	Ala	Val	Gly	Thr	Pro	Asp	Tyr	Leu	Ser	Pro
225				230						235				240	
Glu	Ile	Leu	Gln	Ala	Val	Gly	Gly	Gly	Pro	Gly	Ala	Gly	Ser	Tyr	Gly
			245					250					255		
Pro	Glu	Cys	Asp	Trp	Trp	Ala	Leu	Gly	Val	Phe	Ala	Tyr	Glu	Met	Phe
		260						265				270			
Tyr	Gly	Gln	Thr	Pro	Phe	Tyr	Ala	Asp	Ser	Thr	Ala	Glu	Thr	Tyr	Ala
		275					280					285			

- 31/84 -

Fig. 7 (continued)

Lys	Ile	Val	His	Tyr	Arg	Glu	His	Leu	Ser	Leu	Pro	Leu	Ala	Asp	Thr	290	295	300
Val	Val	Pro	Glu	Glu	Ala	Gln	Asp	Leu	Ile	Arg	Gly	Leu	Leu	Cys	Pro	305	310	315
Ala	Glu	Ile	Arg	Leu	Gly	Arg	Gly	Gly	Ala	Gly	Asp	Phe	Gln	Lys	His	325	330	335
Pro	Phe	Phe	Phe	Gly	Leu	Asp	Trp	Glu	Gly	Leu	Arg	Asp	Ser	Val	Pro	340	345	350
Pro	Phe	Thr	Pro	Asp	Phe	Glu	Gly	Ala	Thr	Asp	Thr	Cys	Asn	Phe	Asp	355	360	365
Val	Val	Glu	Asp	Arg	Leu	Thr	Ala	Met	Val	Ser	Gly	Gly	Gly	Glu	Thr	370	375	380
Leu	Ser	Asp	Met	Gln	Glu	Asp	Met	Pro	Leu	Gly	Val	Arg	Leu	Pro	Phe	385	390	395
Val	Gly	Tyr	Ser	Tyr	Cys	Cys	Met	Ala	Phe	Arg	Asp	Asn	Gln	Val	Pro	405	410	415
Asp	Pro	Thr	Pro	Met	Glu	Leu	Glu	Ala	Leu	Gln	Leu	Pro	Val	Ser	Asp	420	425	430
Leu	Gln	Gly	Leu	Asp	Leu	Gln	Pro	Pro	Val	Ser	Pro	Pro	Asp	Gln	Val	435	440	445
Ala	Glu	Glu	Ala	Asp	Leu	Val	Ala	Val	Pro	Ala	Pro	Val	Ala	Glu	Ala	450	455	460
Glu	Thr	Thr	Val	Thr	Leu	Gln	Gln	Leu	Gln	Glu	Ala	Leu	Glu	Glu	Glu	465	470	475
Val	Leu	Thr	Arg	Gln	Ser	Leu	Ser	Arg	Glu	Leu	Glu	Ala	Ile	Arg	Thr	485	490	495
Ala	Asn	Gln	Asn	Phe	Ser	Ser	Gln	Leu	Gln	Glu	Ala	Glu	Val	Arg	Asn	500	505	510
Arg	Asp	Leu	Glu	Ala	His	Val	Arg	Gln	Leu	Gln	Glu	Arg	Met	Glu	Met	515	520	525
Leu	Gln	Ala	Pro	Gly	Ala	Ala	Ala	Ile	Thr	Gly	Val	Pro	Ser	Pro	Arg	530	535	540
Ala	Thr	Asp	Pro	Pro	Ser	His	Leu	Asp	Gly	Pro	Pro	Ala	Val	Ala	Val	545	550	555
Gly	Gln	Cys	Pro	Leu	Val	Gly	Pro	Gly	Pro	Met	His	Arg	Arg	His	Leu	565	570	575
Leu	Leu	Pro	Ala	Arg	Ile	Pro	Arg	Pro	Gly	Leu	Ser	Glu	Ala	Arg	Cys	580	585	590

Leu Leu Leu Phe Ala Ala Ala Leu Ala Ala Ala Ala Thr Leu Gly Cys
595 600 605
Thr Gly Leu Val Ala Tyr Thr Gly Gly Leu Thr Pro Val Trp Cys Phe
610 615 620
Pro Gly Ala Thr Phe Ala Pro
625 630

- 33/84 -

Fig. 8

BLASTP - alignment of 543_Protein against trembl|AF086824|AF086824_1
 gene: "Crik"; product: "rho/rac-interacting citron kinase"; Mus musculus
 rho/rac-interacting citron kinase (Crik) mRNA, complete cds.
 //:gp|AF086824|3599509 gene: "Crik"; product: "rho/rac-interacting citron
 kinase"; Mus musculus rho/rac-interacting citron kinase (Crik) mRNA,
 complete cds.

This hit is scoring at : 0.0 (expectation value)
 Alignment length (overlap) : 2056
 Identities : 96 %
 Scoring matrix : BLOSUM62 (used to infer consensus pattern)
 Database searched : nrdb_1_;

Q: 1 MLKFKYGARNPLDAGAAEPIASRASRLNLFFQGKPPFMTQQQMSPLSREGILDALFVLFE
 MLKFKYG.RNP :A.A:EPIASRASRLNLFFQGKPP.MTQQQMS.LSREG:LDALF.LFE
 H: 1 MLKFKYGVNRPPEASASEPIASRASRLNLFFQGKPLMTQQQMSALSREGMLDALFALFE

Protein_Kinase_ATP Motif (K binds ATP)

ECSQPALMKIKHVSNFVRKYSDTIAELQELQPSAKDFEVRSLVGC~~GH~~FAEVQVVREKATG
 ECSQPALMK:KHVS:FV:KYSDTIAEL:ELQPSA:DFEVRSLVGC~~GH~~FAEVQVVREKATG
 ECSQPALMKMKHVSSFVQKYSDTIAELRELQPSARDFEVRSLVGC~~GH~~FAEVQVVREKATG

DIYAMKVMKKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKNHLVMEYQPGG
 D:YAMK:MKKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKN:LYLVMEYQPGG
 DVYAMKIMKKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKNNLYLVMEYQPGG

Protein_Kinase_ST Motif (D is an active site)

DLLSLLNRYEDQLDENLIQFYLAELILAVHSVHLMGYVHRDIK~~PENIL~~VDRTGHIKLVDF
 D.LSLLNRYEDQLDE::IQFYLAELILAVSVH MGYVHRDIK~~PENIL~~:DRTG.IKLVDF
 DFLSLLNRYEDQLDESMIQFYLAELILAVSVHQMGYVHRDIK~~PENIL~~IDRTGEIKLVDF

GSAAKMNSNMVNAKLPIGTPDYMAPEVLTVMNGDGKGTGLDCDWWSVGVIAYEM~~IYGR~~
 GSAAKMNSNK V:AKLPIGTPDYMAPEVLTVMN D :GTYGLDCDWWSVGVIAYEM:YG:
 GSAAKMNSNK-VDAKLPIGTPDYMAPEVLTVMNEDRRGTGLDCDWWSVGVIAYEMVY~~GK~~

SPFAEGTSARTFNNIMNFQRFKFPDDPKVSSDFLDLIQSLLCGQKERLKFEGLCCHPFF
 :PF.EGTSARTFNNIMNFQRFKFPDDPKVSS:.LDL:QSLLC QKERLKFEGLCCHPFF
 TPFTEGTSARTFNNIMNFQRFKFPDDPKVSS~~ELLDLL~~QSLLCVQKERLKFEGLCCHPFF

- 34/84 -

Fig. 8 (continued)

SKIDWNNIRNSPPPFVPTLKSDDDTSNFDEPEKNSWVSSSPCQLSPSGFSGEELPFVGFSS
:.DWNINRNSPPPFVPTLKSDDDTSNFDEPEKNSW.P .FSGEELPFVGFSS
ARTDWNINRNSPPPFVPTLKSDDDTSNFDEPEKNSWAFILCVPAEPLAFSGEELPFVGFSS

YSKALGILGRSESVVSGLDSPAKTSSMEKKLLIKSKELQDSQDKCHKMEQEMTRLHRRVS
YSKALG.LGRSESVVS.LDSPAK.SSMEKKLLIKSKELQDSQDKCHKMEQEMTRLHRRVS
YSKALGYLGRSESVVSSLDSPAKVSSMEKKLLIKSKELQDSQDKCHKMEQEMTRLHRRVS

EVEAVLSQKEVELKASETQORSLLQDLATYITECSSLKRSLEQARMEVSQEDDKALQLLH
EVEAVLSQKEVELKASETQORSLLQDLATYITECSSLKRSLEQARMEVSQEDDKALQLLH
EVEAVLSQKEVELKASETQORSLLQDLATYITECSSLKRSLEQARMEVSQEDDKALQLLH

DIREQSRKLQEIKEQEYQAQVEEMRLMMNQLEEDLVSARRRSDLYESELRESRLAAEEFK
DIREQSRKLQEIKEQEYQAQVEEMRLMMNQLEEDLVSARRRSDLYESELRESRLAAEEFK
DIREQSRKLQEIKEQEYQAQVEEMRLMMNQLEEDLVSARRRSDLYESELRESRLAAEEFK

RKATECQHKLKAKDQGKPEVGEYAKLEKINAEQQLKIQELQEKLEKAVKASTEATELLO
RKA.ECQHL:KAKDQGKPEVGEY:KLEKINAEQQLKIQELQEKLEKAVKASTEATELLO
RKANECQHKLKAKDQGKPEVGEYSKLEKINAEQQLKIQELQEKLEKAVKASTEATELLO

NIRQAKERAERELEKLQNREDSSEGIKKLVAAEERRHSLENKVKRLETMERRENRLKDD
NIRQAKERAERELEKL.NREDSSEGI:KKLVAAEERRHSLENKVKRLETMERRENRLKDD
NIRQAKERAERELEKLHNREDSSEGIKKLVAAEERRHSLENKVKRLETMERRENRLKDD

IQTKSQIQQMADKILELEEKHREAQVSAQHLEVHLKQKEQHYYEIKIVLDNQIKKDLAD
IQTKS:QIQQMADKILELEEKHREAQVSAQHLEVHLKQKEQHYYEIKIVLDNQIKKDLAD
IQTKSEQIQQMADKILELEEKHREAQVSAQHLEVHLKQKEQHYYEIKIVLDNQIKKDLAD

KETLENMMQRHEEEAHEKGKILSEQKAMINAMDSKIRSLEQRIVELSEANKLAANSSLFT
KE:LENMMQRHEEEAHEKGKILSEQKAMINAMDSKIRSLEQRIVELSEANKLAANSSLFT
KESLENMMQRHEEEAHEKGKILSEQKAMINAMDSKIRSLEQRIVELSEANKLAANSSLFT

QRNMKAQEEMISELRQOKFYLETQAGKLEAQNRLKEEQLEKISHQDHSDKNRLLLELETRL
QRNMKAQEEMISELRQOKFYLETQAGKLEAQNRLKEEQLEKISHQDHSDK:RLLELETRL
QRNMKAQEEMISELRQOKFYLETQAGKLEAQNRLKEEQLEKISHQDHSDKSRLLELETRL

REVSLEHEEQKLELKRQLTELQLSLQERESQLTALQAARAALESQLRQAKTELEETTAEA
REVSLEHEEQKLELKRQLTELQLSLQERESQLTALQAARAALESQLRQAKTELEETTAEA
REVSLEHEEQKLELKRQLTELQLSLQERESQLTALQAARAALESQLRQAKTELEETTAEA

- 35/84 -

Fig. 8 (continued)

EEEIQALTAHRDEIQRKFDALRNSCTVITDLEEQLNQLTEDNAELNNQNFYLSKQLDEAS
EEEIQALTAHRDEIQRKFDALRNSCTVITDLEEQLNQLTEDNAELNNQNFYLSKQLDEAS
EEEIQALTAHRDEIQRKFDALRNSCTVITDLEEQLNQLTEDNAELNNQNFYLSKQLDEAS

GANDEIVQLRSEVDHLRREITEREMQLTSQKQTMEALKTTCTMLEEQVMDLEALNDELLE
GANDEIVQLRSEVDHLRREITEREMQLTSQKQTMEALKTTCTMLEEQV : DLEALNDELLE
GANDEIVQLRSEVDHLRREITEREMQLTSQKQTMEALKTTCTMLEEQVLDLEALNDELLE

KERQWEAWRSVLGDEKSQFECRVRELQRM LDTEKQSRARADQRITESRQVVELAVKEHKA
KERQWEAWRSVLGDEKSQFECRVRELQRM LDTEKQSRARADQRITESRQVVELAVKEHKA
KERQWEAWRSVLGDEKSQFECRVRELQRM LDTEKQSRARADQRITESRQVVELAVKEHKA

EILALQQALKEQKLKAESLSDKLN DLEKKHAMLEMNARSLOQKLETERELKQRLLEEQAK
EILALQQALKEQKLKAESLSDKLN DLEKKHAMLEMNARSLOQKLETERELKQRLLEEQAK
EILALQQALKEQKLKAESLSDKLN DLEKKHAMLEMNARSLOQKLETERELKQRLLEEQAK

LQQQMDLQKNHIFRLTQGLQEALDRADLLKTERS DLEYQLENIQVLYSHEKVMEGTISQ
LQQQMDLQKNHIFRLTQGLQEALDRADLLKTERS DLEYQLENIQVLYSHEKVMEGTISQ
LQQQMDLQKNHIFRLTQGLQEALDRADLLKTERS DLEYQLENIQVLYSHEKVMEGTISQ

QTKLIDFLQAKMDQPAKKKKVPLQYNELKLALEKEKARCAELEEEALQKTRIELRSAREEA
QTKLIDFLQAKMDQPAKKKKVPLQYNELKLALEKEKARCAELEEEALQKTRIELRSAREEA
QTKLIDFLQAKMDQPAKKKKVPLQYNELKLALEKEKARCAELEEEALQKTRIELRSAREEA

AHRKATDHPHPSTPATARQQIAMS AIVRSPEHQPSAMSL LAPPSSRRKESSTPEEFSRRL
AHRKATDHPHPSTPATARQQIAMS AIVRSPEHQPSAMSL LAPPSSRRKESSTPEEFSRRL
AHRKATDHPHPSTPATARQQIAMS AIVRSPEHQPSAMSL LAPPSSRRKESSTPEEFSRRL

KERMHHNIPHRFNVGLNMRATKCAVCLDTVHFGRQASKCLECQVMCHPKCSTCLPATCGL
KERMHHNIPHRFNVGLNMRATKCAVCLDTVHFGRQASKCLECQVMCHPKCSTCLPATCGL
KERMHHNIPHRFNVGLNMRATKCAVCLDTVHFGRQASKCLECQVMCHPKCSTCLPATCGL

PAEYATHFTEAFCDKMN SPGLQTKEPSSSLHLEGWMKVPRNNKRGQQGWRKYIVLEGS
PAEYATHFTEAFCDKMN SPGLQ : KEP . SSLHLEGWMKVPRNNKRGQQGWRKYIVLEGS
PAEYATHFTEAFCDKMN SPGLQSKEPGSSSLHLEGWMKVPRNNKRGQQGWRKYIVLEGS

KVLIYDNEAREAGQRPVEEFELCLPDGDVSIHGAVGASELANTAKADVPIYILKMESH PHT
KVLIYDNEAREAGQRPVEEFELCLPDGDVSIHGAVGASELANTAKADVPIYILKMESH PHT
KVLIYDNEAREAGQRPVEEFELCLPDGDVSIHGAVGASELANTAKADVPIYILKMESH PHT

- 36/84 -

Fig. 8 (continued)

TCWPGRTLYLLAPSFPDKQRWVTALESVVAGGRVSREKAEADAKLLGNSLLKLEGDDRDL
TCWPGRTLYLLAPSFPDKQRWVTALESVVAGGRVSREKAEADAKLLGNSLLKLEGDDRDL
TCWPGRTLYLLAPSFPDKQRWVTALESVVAGGRVSREKAEADAKLLGNSLLKLEGDDRDL

MNCTLPFSDQVVLVGTEEGLYALNVLKNSLTHVPGIGAVFQIYI IKDLEKLLMIAGEERA
MNCTLPFSDQVVLVGTEEGLYALNVLKNSLTH: PGIGAVFQIYI IKDLEKLLMIAGEERA
MNCTLPFSDQVVLVGTEEGLYALNVLKNSLTHIPGIGAVFQIYI IKDLEKLLMIAGEERA

LCLVDVKVKQSLAQSHLPAQPDISPNI FEAVKGCHLFGAGKIENGLCICAAMPSKVVIL
LCLVDVKVKQSLAQSHLPAQPD: SPNI FEAVKGCHL F. AGKIEN. LCICAAMPSKVVIL
LCLVDVKVKQSLAQSHLPAQPDVSPNI FEAVKGCHLFAAGKIENSLCICAAMPSKVVIL

RYNENLSKYCIRKEIETSEPCSCIHFTNYSILIGTNKFYEIDMKQYTL EFLDKNDHSLA
RYN: NLSKYCIRKEIETSEPCSCIHFTNYSILIGTNKFYEIDMKQYTL: EFLDKNDHSLA
RYNDNLSKYCIRKEIETSEPCSCIHFTNYSILIGTNKFYEIDMKQYTLDEF LDKNDHSLA

PAVFAASSNSFPVSIVQVNSAGQREEYLLCFHEFGVFVDSYGRRSRTDDLKWSRLPLAFA
PAVFA: SSNSFPVSIVQ. NSAGQREEYLLCFHEFGVFVDSYGRRSRTDDLKWSRLPLAFA
PAVFASSNSFPVSIVQANSAGQREEYLLCFHEFGVFVDSYGRRSRTDDLKWSRLPLAFA

YREPYLFVTHFNSLEVIEIQARSSAGTPARAYLDIPNPRYLGPASSGAIYLASSYQDKL
YREPYLFVTHFNSLEVIEIQARSS. G: PARAYL: IPNPRYLGPASSGAIYLASSYQDKL
YREPYLFVTHFNSLEVIEIQARSSLGSPARAYLEIPNPRYLGPASSGAIYLASSYQDKL

RVICCKGNLVKESGTEHHRGPSTSRSSPNKRGPPPTYNEHITKRVASSPAPPEGPSHPREP
RVICCKGNLVKESGTE. HR PSTSRSSPNKRGPPPTYNEHITKRVASSPAPPEGPSHPREP
RVICCKGNLVKESGTEQHRVPSTSRSSPNKRGPPPTYNEHITKRVASSPAPPEGPSHPREP

STPHRY--REGRTELRRDKSPGRPLEREKSPGRMLSTRRERSPGRLFEDSSRGRLPAGAV
STPHRY REGRTELRRDKSPGRPLEREKSPGRMLSTRRERSPGRLFEDSSRGRLPAGAV
STPHRYRDREGRTELRRDKSPGRPLEREKSPGRMLSTRRERSPGRLFEDSSRGRLPAGAV

RTPLSQVNKVWDQSSV 2054

RTPLSQVNKVWDQSSV

RTPLSQVNKVWDQSSV 2055

- 37/84 -

Fig. 9

BLASTP - alignment of 543_Protein against swiss|014578|CTRO_HUMAN
CITRON PROTEIN (FRAGMENT) ./.:treml|AC002563|AC002563_2 gene:
"WUGSC:H_127H14.1";
Human PAC clone 127H14 from 12q, complete sequence. ./.:gp|AC002563|2439517
gene:
"WUGSC:H_127H14.1"; Human PAC clone 127H14 from 12q, complete sequence.

This hit is scoring at : 0.0 (expectation value)
Alignment length (overlap) : 1286
Identities : 100 %
Scoring matrix : BLOSUM62 (used to infer consensus pattern)
Database searched : nrdb_1_;

Q: 769 VLDNQIKKDLADKETLENMMQRHEEEAHEKGKILSEQKAMINAMDSKIRSLEQRIVELSE
VLDNQIKKDLADKETLENMMQRHEEEAHEKGKILSEQKAMINAMDSKIRSLEQRIVELSE
H: 1 VLDNQIKKDLADKETLENMMQRHEEEAHEKGKILSEQKAMINAMDSKIRSLEQRIVELSE

ANKLAANSSSLFTQRNMKAQEEMISELRQOKFYLETQAGKLEAQNRLKEEQLEKISHQDHS
ANKLAANSSSLFTQRNMKAQEEMISELRQOKFYLETQAGKLEAQNRLKEEQLEKISHQDHS
ANKLAANSSSLFTQRNMKAQEEMISELRQOKFYLETQAGKLEAQNRLKEEQLEKISHQDHS

DKNRLLELETRLREVSLEHEEQKLELKRQLTELQLSLQERESQLTALQAARALESQLRQ
DKNRLLELETRLREVSLEHEEQKLELKRQLTELQLSLQERESQLTALQAARALESQLRQ
DKNRLLELETRLREVSLEHEEQKLELKRQLTELQLSLQERESQLTALQAARALESQLRQ

AKTELEETTAAEAEIIIQALTAHRDEIQRKFDALRNSCTVITDLEEQLNQLTEDNAELNNQ
AKTELEETTAAEAEIIIQALTAHRDEIQRKFDALRNSCTVITDLEEQLNQLTEDNAELNNQ
AKTELEETTAAEAEIIIQALTAHRDEIQRKFDALRNSCTVITDLEEQLNQLTEDNAELNNQ

NFYLSKQLDEASGANDEIVQLRSEVDHLRREITEREMQLTSQKQTMEALKTTCTMLEEQV
NFYLSKQLDEASGANDEIVQLRSEVDHLRREITEREMQLTSQKQTMEALKTTCTMLEEQV
NFYLSKQLDEASGANDEIVQLRSEVDHLRREITEREMQLTSQKQTMEALKTTCTMLEEQV

MDLEALNDELLEKERQWEAWRSVLGDEKSQFECRVRELQRMMLDTEKQSRARADQRITESR
MDLEALNDELLEKERQWEAWRSVLGDEKSQFECRVRELQRMMLDTEKQSRARADQRITESR
MDLEALNDELLEKERQWEAWRSVLGDEKSQFECRVRELQRMMLDTEKQSRARADQRITESR

- 38/84 -

Fig. 9 (continued)

QVVELAVKEHKAEILALQQALKEQKLKAESLSDKLNDEKKHAMLEMNARSLQQKLETER
QVVELAVKEHKAEILALQQALKEQKLKAESLSDKLNDEKKHAMLEMNARSLQQKLETER
QVVELAVKEHKAEILALQQALKEQKLKAESLSDKLNDEKKHAMLEMNARSLQQKLETER

ELKQRLLEEQAKLQQQMDLQKNHIFRLTQGLQEALDRADLLKTERS DLEYQLENIQVLYS
ELKQRLLEEQAKLQQQMDLQKNHIFRLTQGLQEALDRADLLKTERS DLEYQLENIQVLYS
ELKQRLLEEQAKLQQQMDLQKNHIFRLTQGLQEALDRADLLKTERS DLEYQLENIQVLYS

HEKVKMEGTISQQTKLIDFLQAKMDQPAKKKKVPLQYNELKLALEKEKARCAELEEEALQK
HEKVKMEGTISQQTKLIDFLQAKMDQPAKKKKVPLQYNELKLALEKEKARCAELEEEALQK
HEKVKMEGTISQQTKLIDFLQAKMDQPAKKKKVPLQYNELKLALEKEKARCAELEEEALQK

TRIELRSAREEAAHRKATDHPHPSTPATARQQIAMS AIVRSPEHQPSAMSLAPPSSRRK
TRIELRSAREEAAHRKATDHPHPSTPATARQQIAMS AIVRSPEHQPSAMSLAPPSSRRK
TRIELRSAREEAAHRKATDHPHPSTPATARQQIAMS AIVRSPEHQPSAMSLAPPSSRRK

ESSTPEEFSRRLKERMHHNIPHRFNVGLNMRATKCAVCLDTVHFGRQASKCLECQVMCHP
ESSTPEEFSRRLKERMHHNIPHRFNVGLNMRATKCAVCLDTVHFGRQASKCLECQVMCHP
ESSTPEEFSRRLKERMHHNIPHRFNVGLNMRATKCAVCLDTVHFGRQASKCLECQVMCHP

KCSTCLPATCGLPAEYATHFTEAFCDKMNSPGLQTKESPSSSLHLEGWMKVPRNNKRGQQ
KCSTCLPATCGLPAEYATHFTEAFCDKMNSPGLQTKESPSSSLHLEGWMKVPRNNKRGQQ
KCSTCLPATCGLPAEYATHFTEAFCDKMNSPGLQTKESPSSSLHLEGWMKVPRNNKRGQQ

GWDRKYIVLEGSKVLIYDNEAREAGQRPVEEFELCLPDGDVSIHGAVGASELANTAKADV
GWDRKYIVLEGSKVLIYDNEAREAGQRPVEEFELCLPDGDVSIHGAVGASELANTAKADV
GWDRKYIVLEGSKVLIYDNEAREAGQRPVEEFELCLPDGDVSIHGAVGASELANTAKADV

PYILKMESHPTHCTCWPGR TLYLLAPSFDPKQRWVTALESV VAGGRVSREKAEADAKLLGN
PYILKMESHPTHCTCWPGR TLYLLAPSFDPKQRWVTALESV VAGGRVSREKAEADAKLLGN
PYILKMESHPTHCTCWPGR TLYLLAPSFDPKQRWVTALESV VAGGRVSREKAEADAKLLGN

SLLKLEGDDRLDMNCTLPFSDQVVLVGTEEGLYALNVLKNSLTHVPGIGAVFQIYI IKDL
SLLKLEGDDRLDMNCTLPFSDQVVLVGTEEGLYALNVLKNSLTHVPGIGAVFQIYI IKDL
SLLKLEGDDRLDMNCTLPFSDQVVLVGTEEGLYALNVLKNSLTHVPGIGAVFQIYI IKDL

EKLLMIAGEERALCLVDVKKVKQSLAQSHLPAQPDISPNI FEAVKGCHLFGAGKIENGLC
EKLLMIAGEERALCLVDVKKVKQSLAQSHLPAQPDISPNI FEAVKGCHLFGAGKIENGLC
EKLLMIAGEERALCLVDVKKVKQSLAQSHLPAQPDISPNI FEAVKGCHLFGAGKIENGLC

- 39/84 -

Fig. 9 (continued)

ICAAMPSKVILRYNENLSKYCIRKEIETSEPCSCIHFTNYSILIGTNKFYEIDMKQYTL
ICAAMPSKVILRYNENLSKYCIRKEIETSEPCSCIHFTNYSILIGTNKFYEIDMKQYTL
ICAAMPSKVILRYNENLSKYCIRKEIETSEPCSCIHFTNYSILIGTNKFYEIDMKQYTL

EEFLDKNDHSLAPAVFAASSNSFPVSIVQVNSAGQREEYLLCFHEFGVFVDSYGRRSRTD
EEFLDKNDHSLAPAVFAASSNSFPVSIVQVNSAGQREEYLLCFHEFGVFVDSYGRRSRTD
EEFLDKNDHSLAPAVFAASSNSFPVSIVQVNSAGQREEYLLCFHEFGVFVDSYGRRSRTD

DLKWSRLPLAFAYREPYLEFVTHFNSLEVIEIQARSSAGTPARAYLDIPNPRYLGPASSG
DLKWSRLPLAFAYREPYLEFVTHFNSLEVIEIQARSSAGTPARAYLDIPNPRYLGPASSG
DLKWSRLPLAFAYREPYLEFVTHFNSLEVIEIQARSSAGTPARAYLDIPNPRYLGPASSG

AIYLASSYQDKLRVICCKGNLVKESGTEHHRGPSTSRSSPNKRGPPPTYNEHITKRVASSP
AIYLASSYQDKLRVICCKGNLVKESGTEHHRGPSTSRSSPNKRGPPPTYNEHITKRVASSP
AIYLASSYQDKLRVICCKGNLVKESGTEHHRGPSTSRSSPNKRGPPPTYNEHITKRVASSP

APPEGPSHPREPSTPHRYREGRTTELRRDKSPGRPLEREKSPGRMLSTRRERSPGRLFEDS
APPEGPSHPREPSTPHRYREGRTTELRRDKSPGRPLEREKSPGRMLSTRRERSPGRLFEDS
APPEGPSHPREPSTPHRYREGRTTELRRDKSPGRPLEREKSPGRMLSTRRERSPGRLFEDS

SRGRLPAGAVRTPLSQVNKVWDQSSV 2054
SRGRLPAGAVRTPLSQVNKVWDQSSV
SRGRLPAGAVRTPLSQVNKVWDQSSV 1286

- 40/84 -

Fig. 10

BLASTP - alignment of 543_Protein against aageneseq|AAB43359|AAB43359
Human ORFX ORF3123 polypeptide sequence SEQ ID NO:6246.

This hit is scoring at : 0.0 (expectation value)
Alignment length (overlap) : 1286
Identities : 100 %
Scoring matrix : BLOSUM62 (used to infer consensus pattern)
Database searched : aageneseq

Q: 769 VLDNQIKKDLADKETLENMMQRHEEEAHEKGKILSEQKAMINAMDSKIRSLEQRIVELSE
VLDNQIKKDLADKETLENMMQRHEEEAHEKGKILSEQKAMINAMDSKIRSLEQRIVELSE
H: 1 VLDNQIKKDLADKETLENMMQRHEEEAHEKGKILSEQKAMINAMDSKIRSLEQRIVELSE

ANKLAANSSSLFTQRMKAQEEMISELRQKFYLETQAGKLEAQNRLKEEQLEKISHQDHS
ANKLAANSSSLFTQRMKAQEEMISELRQKFYLETQAGKLEAQNRLKEEQLEKISHQDHS
ANKLAANSSSLFTQRMKAQEEMISELRQKFYLETQAGKLEAQNRLKEEQLEKISHQDHS

DKNRLLLELETRLREVSLEHEEQKLELKRQLTELQLSLQERESQLTALQAARALESQLRQ
DKNRLLLELETRLREVSLEHEEQKLELKRQLTELQLSLQERESQLTALQAARALESQLRQ
DKNRLLLELETRLREVSLEHEEQKLELKRQLTELQLSLQERESQLTALQAARALESQLRQ

AKTELEETTAAAEIIIQALTAHRDEIQRKFDALRNSCTVITDLEEQLNQLTEDNAELNNQ
AKTELEETTAAAEIIIQALTAHRDEIQRKFDALRNSCTVITDLEEQLNQLTEDNAELNNQ
AKTELEETTAAAEIIIQALTAHRDEIQRKFDALRNSCTVITDLEEQLNQLTEDNAELNNQ

NFYLSKQLDEASGANDEIVQLRSEVDHLRREITEREMQLTSQKQTMEALKTTCTMLEEQV
NFYLSKQLDEASGANDEIVQLRSEVDHLRREITEREMQLTSQKQTMEALKTTCTMLEEQV
NFYLSKQLDEASGANDEIVQLRSEVDHLRREITEREMQLTSQKQTMEALKTTCTMLEEQV

MDLEALNDELLEKERQWEAWRSVLGDEKSQFECRVRELQRM LDTEKQSRARADQRITESR
MDLEALNDELLEKERQWEAWRSVLGDEKSQFECRVRELQRM LDTEKQSRARADQRITESR
MDLEALNDELLEKERQWEAWRSVLGDEKSQFECRVRELQRM LDTEKQSRARADQRITESR

QVVELAVKEHKAEILALQQALKEQKLKAESLSDKLNDEKKHAMLEMNARSLQQKLETER
QVVELAVKEHKAEILALQQALKEQKLKAESLSDKLNDEKKHAMLEMNARSLQQKLETER
QVVELAVKEHKAEILALQQALKEQKLKAESLSDKLNDEKKHAMLEMNARSLQQKLETER

- 41/84 -

Fig. 10 (continued)

ELKQRLLEEQAQLQQQMDLQKNHIFRLTQGLQEALDRADLLKTERS DLEYQLENIQVLYS
ELKQRLLEEQAQLQQQMDLQKNHIFRLTQGLQEALDRADLLKTERS DLEYQLENIQVLYS
ELKQRLLEEQAQLQQQMDLQKNHIFRLTQGLQEALDRADLLKTERS DLEYQLENIQVLYS

HEKVKMEGTISQOTKLIDFLQAKMDQPAKKKKVPLQYNELKLAL EKEKARCABLEEALQK
HEKVKMEGTISQOTKLIDFLQAKMDQPAKKKKVPLQYNELKLAL EKEKARCABLEEALQK
HEKVKMEGTISQOTKLIDFLQAKMDQPAKKKKVPLQYNELKLAL EKEKARCABLEEALQK

TRIELRSAREEAAHRKATDHPHPSTPATARQQIAMS AIVRSPEHQPSAMSL LAPPSSRRK
TRIELRSAREEAAHRKATDHPHPSTPATARQQIAMS AIVRSPEHQPSAMSL LAPPSSRRK
TRIELRSAREEAAHRKATDHPHPSTPATARQQIAMS AIVRSPEHQPSAMSL LAPPSSRRK

ESSTPEEFSRRLKERMHHNIPHRFNVGLNMRATKCAVCLD TVHFGRQASKCLECQVMCHP
ESSTPEEFSRRLKERMHHNIPHRFNVGLNMRATKCAVCLD TVHFGRQASKCLECQVMCHP
ESSTPEEFSRRLKERMHHNIPHRFNVGLNMRATKCAVCLD TVHFGRQASKCLECQVMCHP

KCSTCLPATCGLPAEYATHFTEAFCRDKMNSPGLQTKEPSSSLHLEGWMKVPRNNKRGQQ
KCSTCLPATCGLPAEYATHFTEAFCRDKMNSPGLQTKEPSSSLHLEGWMKVPRNNKRGQQ
KCSTCLPATCGLPAEYATHFTEAFCRDKMNSPGLQTKEPSSSLHLEGWMKVPRNNKRGQQ

GWDRKYIVLEGSKVLIYDNEAREAGQRPVEEFELCLPDGDVS IHGAVGASELANTAKADV
GWDRKYIVLEGSKVLIYDNEAREAGQRPVEEFELCLPDGDVS IHGAVGASELANTAKADV
GWDRKYIVLEGSKVLIYDNEAREAGQRPVEEFELCLPDGDVS IHGAVGASELANTAKADV

PYILKMESHPTHTCWPGR TLYLLAPSFDPKQRWVTALESVVAGGRVSREKAEADAKLLGN
PYILKMESHPTHTCWPGR TLYLLAPSFDPKQRWVTALESVVAGGRVSREKAEADAKLLGN
PYILKMESHPTHTCWPGR TLYLLAPSFDPKQRWVTALESVVAGGRVSREKAEADAKLLGN

SLLKLEGDDRLDMNCTLPFSDQVVLVGTEEGLYALNVLKNSLTHVPGIGAVFQIYIIKDL
SLLKLEGDDRLDMNCTLPFSDQVVLVGTEEGLYALNVLKNSLTHVPGIGAVFQIYIIKDL
SLLKLEGDDRLDMNCTLPFSDQVVLVGTEEGLYALNVLKNSLTHVPGIGAVFQIYIIKDL

EKLLMIAGEERALCLVDVKKVKQSLAQSHLPAQPDISP NIFEAVKGCHLFGAGKIENGLC
EKLLMIAGEERALCLVDVKKVKQSLAQSHLPAQPDISP NIFEAVKGCHLFGAGKIENGLC
EKLLMIAGEERALCLVDVKKVKQSLAQSHLPAQPDISP NIFEAVKGCHLFGAGKIENGLC

ICAAMPSKVILRYNENLSKYCIRKEIETSEPCSCIHFTNYSILIGTNKFYEIDMKQYTL
ICAAMPSKVILRYNENLSKYCIRKEIETSEPCSCIHFTNYSILIGTNKFYEIDMKQYTL
ICAAMPSKVILRYNENLSKYCIRKEIETSEPCSCIHFTNYSILIGTNKFYEIDMKQYTL

- 42/84 -

Fig. 10 (continued)

EEFLDKNDHSLAPAVFAASSNSFPVSIVQVNSAGQREEYLLCFHEFGVFVDSYGRRSRTD
EEFLDKNDHSLAPAVFAASSNSFPVSIVQVNSAGQREEYLLCFHEFGVFVDSYGRRSRTD
EEFLDKNDHSLAPAVFAASSNSFPVSIVQVNSAGQREEYLLCFHEFGVFVDSYGRRSRTD

DLKWSRLPLAFAYREPYLFVTHFNSLEVIEIQARSSAGTPARAYLDIPNPRYLGPASSG
DLKWSRLPLAFAYREPYLFVTHFNSLEVIEIQARSSAGTPARAYLDIPNPRYLGPASSG
DLKWSRLPLAFAYREPYLFVTHFNSLEVIEIQARSSAGTPARAYLDIPNPRYLGPASSG

AIYLASSYQDKLRVICCKGNLVKESGTEHHRGPSTSRSSPNKRGPPPTYNEHITKRVASSP
AIYLASSYQDKLRVICCKGNLVKESGTEHHRGPSTSRSSPNKRGPPPTYNEHITKRVASSP
AIYLASSYQDKLRVICCKGNLVKESGTEHHRGPSTSRSSPNKRGPPPTYNEHITKRVASSP

APPEGPSHPREPSTPHRYREGRTTELRRDKSPGRPLEREKSPGRMLSTRRERSPGRLFEDS
APPEGPSHPREPSTPHRYREGRTTELRRDKSPGRPLEREKSPGRMLSTRRERSPGRLFEDS
APPEGPSHPREPSTPHRYREGRTTELRRDKSPGRPLEREKSPGRMLSTRRERSPGRLFEDS

SRGRLPAGAVRTPLSQVNKVWDQSSV 2054
SRGRLPAGAVRTPLSQVNKVWDQSSV
SRGRLPAGAVRTPLSQVNKVWDQSSV 1286

- 43/84 -

Fig. 11

BLASTP - alignment of 543_Protein against trembl|AB023166|AB023166_1
gene: "KIAA0949"; product: "KIAA0949 protein"; Homo sapiens mRNA for
KIAA0949
protein, partial cds. //:gp|AB023166|4589542 gene: "KIAA0949"; product:
"KIAA0949 protein"; Homo sapiens mRNA for KIAA0949 protein, partial cds.

This hit is scoring at : 0.0 (expectation value)

Alignment length (overlap) : 940

Identities : 100 %

Scoring matrix : BLOSUM62 (used to infer consensus pattern)

Database searched : nrdb_1_;

Q: 1115 QSRARADQRITESRQVVELAVKEHKAEILALQQALKEQKLKAESLSDKLNDEKKHAMLE
QSRARADQRITESRQVVELAVKEHKAEILALQQALKEQKLKAESLSDKLNDEKKHAMLE
H: 1 QSRARADQRITESRQVVELAVKEHKAEILALQQALKEQKLKAESLSDKLNDEKKHAMLE

MNARSLQQKLETERELKQRLLEEQAKLQQQMDLQKNHIFRLTQGLQEALDRADLLKTERS
MNARSLQQKLETERELKQRLLEEQAKLQQQMDLQKNHIFRLTQGLQEALDRADLLKTERS
MNARSLQQKLETERELKQRLLEEQAKLQQQMDLQKNHIFRLTQGLQEALDRADLLKTERS

DLEYQLENIQVLYSHEKVMEGTISQQTKLIDFLQAKMDQPAKKKKVPLQYNELKLALAEK
DLEYQLENIQVLYSHEKVMEGTISQQTKLIDFLQAKMDQPAKKKKVPLQYNELKLALAEK
DLEYQLENIQVLYSHEKVMEGTISQQTKLIDFLQAKMDQPAKKKKVPLQYNELKLALAEK

EKARCAELEEEALQKTRIELRSAREEAAHRKATDHPHPSTPATARQQIAMSIVRSPEHQF
EKARCAELEEEALQKTRIELRSAREEAAHRKATDHPHPSTPATARQQIAMSIVRSPEHQF
EKARCAELEEEALQKTRIELRSAREEAAHRKATDHPHPSTPATARQQIAMSIVRSPEHQF

SAMSLAPPSRRKESSTPEEFSRRLKERMHHNIPHRFNVGLNMRATKCAVCLDTVHFGR
SAMSLAPPSRRKESSTPEEFSRRLKERMHHNIPHRFNVGLNMRATKCAVCLDTVHFGR
SAMSLAPPSRRKESSTPEEFSRRLKERMHHNIPHRFNVGLNMRATKCAVCLDTVHFGR

QASKCLECQVMCHPKCSTCLPATCGLPAEYATHFTEAFCDKMNSPGLQTKEPSSSLHLE
QASKCLECQVMCHPKCSTCLPATCGLPAEYATHFTEAFCDKMNSPGLQTKEPSSSLHLE
QASKCLECQVMCHPKCSTCLPATCGLPAEYATHFTEAFCDKMNSPGLQTKEPSSSLHLE

GWMKVPRNNKRGQQGWRKYIVLEGSKVLIYDNEAREAGQRPVEEFELCLPDGDVSIHGA
GWMKVPRNNKRGQQGWRKYIVLEGSKVLIYDNEAREAGQRPVEEFELCLPDGDVSIHGA
GWMKVPRNNKRGQQGWRKYIVLEGSKVLIYDNEAREAGQRPVEEFELCLPDGDVSIHGA

- 44/84 -

Fig 11 (continued)

VGASELANTAKADVPIILKMESHPTTCWPGRPLYLLAPSFDPKQRWVTALESVVAGGRV
VGASELANTAKADVPIILKMESHPTTCWPGRPLYLLAPSFDPKQRWVTALESVVAGGRV
VGASELANTAKADVPIILKMESHPTTCWPGRPLYLLAPSFDPKQRWVTALESVVAGGRV

SREKAEADAKLLGNSLLKLEGDDRLDMNCTLPFSDQVVLVGTEEGLYALNVLKNSLTHVP
SREKAEADAKLLGNSLLKLEGDDRLDMNCTLPFSDQVVLVGTEEGLYALNVLKNSLTHVP
SREKAEADAKLLGNSLLKLEGDDRLDMNCTLPFSDQVVLVGTEEGLYALNVLKNSLTHVP

GIGAVFQIYIIKDLEKLLMIAGEERALCLVDVKKVKQSLAQSHLPAQPDISPNI FEAVKG
GIGAVFQIYIIKDLEKLLMIAGEERALCLVDVKKVKQSLAQSHLPAQPDISPNI FEAVKG
GIGAVFQIYIIKDLEKLLMIAGEERALCLVDVKKVKQSLAQSHLPAQPDISPNI FEAVKG

CHLFGAGKIENGLCICAAMPSKVILRYNENLSKYCIRKEIETSEPCSCIHFTNYSILIG
CHLFGAGKIENGLCICAAMPSKVILRYNENLSKYCIRKEIETSEPCSCIHFTNYSILIG
CHLFGAGKIENGLCICAAMPSKVILRYNENLSKYCIRKEIETSEPCSCIHFTNYSILIG

TNKFYEIDMKQYTL EEFLDKNDHSLAPAVFAASSNSFPVSIVQVNSAGQREEYLLCFHEF
TNKFYEIDMKQYTL EEFLDKNDHSLAPAVFAASSNSFPVSIVQVNSAGQREEYLLCFHEF
TNKFYEIDMKQYTL EEFLDKNDHSLAPAVFAASSNSFPVSIVQVNSAGQREEYLLCFHEF

GVFVDSYGRRSRTDDLKWSRLPLAFAYREPYLFVTHFNSLEVIEIQARSSAGTPARAYLD
GVFVDSYGRRSRTDDLKWSRLPLAFAYREPYLFVTHFNSLEVIEIQARSSAGTPARAYLD
GVFVDSYGRRSRTDDLKWSRLPLAFAYREPYLFVTHFNSLEVIEIQARSSAGTPARAYLD

IPNPRYLGPASSGAIYLASSYQDKLRVICCKGNLVKESGTEHHRGPSTSRSSPNKRGP
IPNPRYLGPASSGAIYLASSYQDKLRVICCKGNLVKESGTEHHRGPSTSRSSPNKRGP
IPNPRYLGPASSGAIYLASSYQDKLRVICCKGNLVKESGTEHHRGPSTSRSSPNKRGP

TYNEHITKRVASSPAPPEGPSHPREPSTPHRYREGRTTELRRDKSPGRPLEREKSPGRMLS
TYNEHITKRVASSPAPPEGPSHPREPSTPHRYREGRTTELRRDKSPGRPLEREKSPGRMLS
TYNEHITKRVASSPAPPEGPSHPREPSTPHRYREGRTTELRRDKSPGRPLEREKSPGRMLS

TRRERSPGRLFEDSSRGRLPAGAVRTPLSQVNKVWDQSSV 2054
TRRERSPGRLFEDSSRGRLPAGAVRTPLSQVNKVWDQSSV
TRRERSPGRLFEDSSRGRLPAGAVRTPLSQVNKVWDQSSV 940

- 45/84 -

Fig. 12

BLASTP - alignment of 543_Protein against swissnew|P54265|DMK_MOUSE
 MYOTONIN-PROTEIN KINASE (EC 2.7.1.-) (MYOTONIC DYSTROPHY PROTEIN KINASE) (MDPK)
 (DM-KINASE) (DMK) (DMPK) (MT-PK).//:swiss|P54265|DMK_MOUSE MYOTONIN-PROTEIN
 KINASE

(EC 2.7.1.-) (MYOTONIC DYSTROPHY PROTEIN KINASE) (MDPK) (DM-KINASE) (DMK) (DMPK)
 (MT-PK).//:trernbl|Z38015|MMMDMPK_1 gene: "DM-PK"; product: "myotonic dystrophy
 protein kinase"; M.musculus DMR-N9 gene, exons 4 and 5, and DM-PK gene encoding
 myotonic dystrophy protein kinase //:gp|Z38015|563526 gene: "DM-PK"; product:
 "myotonic dystrophy protein kinase"; M.musculus DMR-N9 gene, exons 4 and 5, and
 DM-PK gene encoding myotonic dystrophy protein kinase.

This hit is scoring at : 3e-89 (expectation value)

Alignment length (overlap) : 522

Identities : 38 %

Scoring matrix : BLOSUM62 (used to infer consensus pattern)

Database searched : nrdb_1_;

Q: 46 LSREGILDALFVLFEECSPALMKIKHVSNFVRKYSDTIAELQELQPSAKDFEVRSLVGC
 L. E :LD.L. ::E.... L.: K:V::F::A.L:E:: ...DFE: :::G
 H: 20 LGLEPLDLLLLGVHQELGASHLAQDKYVADFLQWVEPIAARLKEVRLQRDDFEILKVIGR

GHFAEVQVVREKATGDIYAMKVMKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQ
 G F:EV.VV: K.TG.:YAMK:M.K :L.: :VS F.EER::L WI.QL.:AFQ
 GAFSEVAVVMKQGTGQVYAMKIMNKWDMLKRGEVSCFREERDVLVKGDRRWITQLHFAFQ

DKNHLYLVMEYQPGGDLSSLNRYEDQLDENLIQFYLAELILAVHSVHLMGYVHRDIKPE
 D:N:LYLVMEY. GGDLL:LL::: ::::.....:FYLAE:::A:.SVH :GYVHRDIKPE:
 DENYLYLVMEYVGGDLLTLLSKFGERIPAEMARFYLAELVMAIDSVHRLGYVHRDIKPD

NILVDRTGHIKLVDFGSAKMNSNMVNAKLPIGTPDYMAPEVL-TVMNGDGKGTGYGLDC
 NIL:DR.GHI:L.DFGS..K:... MV.: ::GTPDY::PE:L .V .G.G.G:YG :C
 NILLDRCGHIRLADFGSCLKLQPDGMVRSVLVAVGTPDYLSPEILQAVGGGPGAGSYGPEC

DWWSVGVIAYEMIYGRSPFAEGTSARTFNNIMNFQRFLKFP-DDPKVSSDFLDLIQSLLC
 DWW::GV.AYEM.YG::PFA.T: .I:::..L..P D. V... DLI::LLC
 DWWALGVFAYEMFYGQTPFYADSTAETYAKIVHYREHLSLPLADTVVPPEAQDLIRGLLC

- 46/84 -

Fig. 12 (continued)

GQKERLKFEG---LCCHPFFSKIDWNNIRNSPPPFVPTLKSDDDTSNFD--EPEKNSWVS
.: RL G . HPFF :DW.:R:S PPF.P.... .DT.NFD E.. :.:VS
PAEIRLGRGGAGDFQKHPFFFGLDWEGLRDSVPPFTPDFEGATDTCNFDVVEDRLTAMVS

SSPCQLSPSGFS---GEELPFVGFYSYKALGILGRSESVVSGLDSPAKTSSMEKKLLIKS
.. .LS.. . G .LPFVG:SY . :. R...V P .T.. :.L :.
GGGETLSDMQEDMPLGVRLPFVGYSY---CCMAFRDNQV-----PDPTMELEALQLPV

KELQDSQDKCHKMEQEMTRLHRRVSEVEAVLSQKEVELKASETQRSILLEQDLATYITECS
.:LQ. . : : .V.A :.:E.Q.:L E: L.. .
SDLQGLDLQPPVSPPDQVAEEADLVAVPAPVAEATTVTTLQQLQEALEEEVLTR-----Q

SLKRSLE---QARMEVSQEDDKALQLLHDIREQSRKLQEIKE 554
SL.R.LE .A....S.: :A .D:... R:LQE .E
SLSRELEAIRTANQNFSSQLQEAEVRNRDLEAHVRQLQERME 527

- 47/84 -

Fig. 13

BLASTP - alignment of 543_Protein against pdb|1CDK|1CDK-A
 camp-dependent protein kinase(protein kinase a)protein kinase
 inhibitor(pki(5-24))

This hit is scoring at : 4e-44 (expectation value)

Alignment length (overlap) : 333

Identities : 33 %

Scoring matrix : BLOSUM62 (used to infer consensus pattern)

Database searched : nrdb_1_;

Q: 71 KHVSNFVRKYSDTIAELQELQPSAKDFEVRSLVGCCHFAEVQVVREKATGDIYAMKVMKK
 K ..F::K::.....L. .FE ...G.G.F..V.:V.:K.TG: :AMK::K
 H: 14 KAKEDFLKKWENPAQNTAHLD----QFERIKTLGTGSFGRVMLVKHKETGNHFAMKILDK

KALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKNHLVLMVEYQPGGDLSSLNRYE
 :::: :Q:: .E::IL.. .P:: :L:Y:F:D::LY:VMEY PGG::S L.R.
 QKVVKLKQIEHTLNEKRILQAVNFPFLVKLEYSFKDNSNLYMVMEYVPGGEMFSLRRI-

DQLDENLIQFYLAELILAVHSVHLMGYVHRDIKPENILVDRTGHIKLVDFGSAKMNSNK
E ..FY.A::L... :H :...:RD:KPEN:L:D:.G:I::DFG A:::....
 GRFSEPHARFYAAQIVLTFEYLSLDLIYRDLKPENLLIDQQGYIQVTDGFAKRVKGR

MVNAKLPIGTPDYMAPEVLTVMNGDGKGTYGLDCDWWSVGVIAYEMIYGRSPFAEGTSAR
 ... GTP:Y:APE:: .KG Y. .DW::GV:.YEM. G .PF
 WTLG----GTPEYLAPEIIL-----SKG-YNKAVDWWALGVLIYEMAAGYPPFFADQPIQ

TFNNIMNFQRFKFPDDPKVSSDFLDLIQSLLCGQKERLKFEG-----LCCHPFFSK
I:: : ::FP. ..SSD. DL:::LL Q : .K G : H::F::
 IYEKIVSGK--VRFPS--HFSSDLKDLLRNLL--QVDLTKRFGNLKDGVNNDIKNHKWFAT

IDWNNI--RNSPPPFVPTLKSDDDTSNFDEPEK 393

.DW I R. ..PF:P..K...DTSNFD: E:

TDWIAIYQKVEAPFIPKFKGPGDTSNFDYEE 325

This hit is scoring at : 219.4 E=5.5e-62
Scoring matrix : BLOSUM62 (used to infer consensus pattern)

```

-----GTSARTFNIMNfqgrflKFPDDPKVSSDFLDLIQSLLC-GQKERL---KFEGLCCH
      .. ...F.  ..      :P ....S.:. DL:..L  ...:R      . : :. H
ktridpleelfrikkr.....rlplpsncSeelkdLlkkcLnkDPskRpGsatakeilnh

```

PFF	360
P:F	
pwf	278

- 49/84 -

Fig. 15

HMMPFAM - alignment of 543_Protein against pfam|hmm|PH
PH domain

This hit is scoring at : 45.8 E=1.8e-11

Scoring matrix : BLOSUM62 (used to infer consensus pattern)

Q: 1471 LHLEGWMKVPRnnkrgqQGWD RKYIVLEGSKVLIYDNE-AREAGQRPVEEFELCLpDGdv
: EGW: .. :.W.:Y.VL :. :L.Y.:.G. P:. :. :.
H: 1 vikeGwLlkks.....kswkkRyfvLfnnvLlyykdskkkpksipLsgcqvek.pd..

sihgavgaselantakadvPYILKMESHpHttcwpgrTLYLLAPSFDPKQRWVTALESVV
.....: .. TL.L A.SWV.A::S:
.....kncFeirt.dr.....tlllqaeseeerkeWvkaiqsai

A 1590

.

r 85

- 50/84 -

Fig. 16

HMMPFAM - alignment of 543_Protein against pfam|hmm|CNH
 CNH domain (Domain found in NIK1-like kinase, mouse citron and yeast ROM1,
 ROM2)

This hit is scoring at : 380.7 1.5e-110

Scoring matrix : BLOSUM62 (used to infer consensus pattern)

```

Q: 1619 LDMNCTLPFSDQ-----VVLVGTEEGLYALNVLKN-----SLTHVPGIGAVFQIYII
      ....C. P.: .      :LVGTEEGLY.LN: ..      :L... :V QI:::
H: 1 ytakcnhpitcdaLWGkiLLvgTeeGLYvlnisdqlnkdhfegtlekiisrrsvtqiwvl

      KDLEKLLMIAGE---ERALCLVDVKV-----KQSLAQSHLPAQPDISPNIFFEA
      :: . LLMI:G:      A L :: :      K.:L...L ..... . .E
      eennvLlmisGkkpylyahpLsglvekklaqknspisikdalgsarlviRknvlsvkied

      VKG--CHLFGA-GKIEN--GLCICAAMPSKVIL--RYNENLSKYCIR-----
      VKG CHLF.. . :. L :.AA:.S.V :L YN. . ...
      vkGNSchlfavkvngkrakilflaaalkssvqllaqwynplkkfklfkssNNiLNNEled

      -----KEIETSEPCSCIHFTNY---SILIGTNK----FYEIDMKQYTL-----E
      K I . . . : : : . I . IG.:K . . D: Q: .
      ikkfklklivpvppllvveltsssfelpkiciGvdkPVGgeagfdvvqfhqtphlNslkfks

      EFLDKNDHSLAPAVFAASS-----NSFPVSIVQVNSAGQREEYLLCFHEFGVFVDSYG--
      ....K.D SL. A: ..S. ....V IVQ :..GQR:E.LLCF.EFGVVFV: .G
      slvskedlslpnaleetskkiaTCkpisviivqgsdgGqRdelLLcfdefgvfVNlqGae

      -RRSRTDDLKWSRLPLAFAYREPYLFVTHFNSLEVIEIQARSS-AGTPARA-YLDIPNPR
      RRSR.. L.W. :P AFAY EPYL.. H N.:E: EI:. . . . A . .L:. . R
      arrsrkpiltwefmpeafayvepyllafhsngieIreietgelNlqeladralllearkir

      YLGP-AISSGAIYLASSY      1916
      .LG. .IS. .I.L:SS
      lLgsCeisdrkIllssp      378
  
```

- 51/84 -

Fig. 17

HMMPFAM - alignment of 543_Protein against pfam|hmm|DAG_PE-bind
Phorbol esters/diacylglycerol binding

This hit is scoring at : 28.7 E=6.1e-05

Scoring matrix : BLOSUM62 (used to infer consensus pattern)

```
Q: 1390 HRFNVGLN-MRATKCAVCLDTVHFG-RQASKCLECQVMCHPKCSTCLPATC 1438
      HRF. . . .T C C :.: :Q. KC .C : .H.:C.. :P..C
H: 1 HrFkrtrtfyksptfCdHcGellwglakQGlkCsnCglnvHkrChekVptnC 51
```

Note: Phorbol esters/diacylglycerol binding domain also as the Protein kinase C conserved region 1 (C1) domain. Diacylglycerol (DAG) is an important second messenger. Phorbol esters (PE) are analogues of DAG and potent tumor promoters that cause a variety of physiological changes when administered to both cells and tissues. DAG activates a family of serine/threonine protein kinases, collectively known as protein kinase C (PKC). Phorbol esters can directly stimulate PKC. The N-terminal region of PKC, known as C1, has been shown to bind PE and DAG in a phospholipid and zinc-dependent fashion. The C1 region contains one or two copies (depending on the isozyme of PKC) of a cysteine-rich domain about 50 amino-acid residues long and essential for DAG/PE-binding.

- 52/84 -

Fig. 18

HMMPFAM - alignment of 543_Protein against pfam|hmm|pkinase_C
Protein kinase C terminal domain

This hit is scoring at : 15.4 E=0.0018

Scoring matrix : BLOSUM62 (used to infer consensus pattern)

```
Q: 361 SKIDWNNI--RNSPPPFVPTLKSDDDTSNFDE 390
    ..IDW::: :. .PPF P.:KS. DTSNFD:
H: 1 reIdWdkLEnkeiePPFKPkiksprDtsNFDk 32
```

- 53/84 -

Fig. 19

Prosites results:

PS00479	1390->1439	DAG_PE_BIND_DOM_1	PDOC00379
PS00029	854->876	LEUCINE_ZIPPER	PDOC00029
PS00029	991->1013	LEUCINE_ZIPPER	PDOC00029
PS00029	1057->1079	LEUCINE_ZIPPER	PDOC00029
PS00029	1159->1181	LEUCINE_ZIPPER	PDOC00029
PS00107	103->127	PROTEIN_KINASE_ATP	PDOC00100
PS00108	217->230	PROTEIN_KINASE_ST	PDOC00100
PS00867	1172->1180	CPSASE_2	PDOC00676

- 54/84 -

Fig. 20

genewise output:

```

gi|3599509|gb|A      1 MLKFKYGVRNPPEASASEPIASRASRLNLFFQ
                      MLKFKYG RNP +A A+EPIASRASRLNLFFQ
                      MLKFKYGARNPLDAGAAEPIASRASRLNLFFQ
gi|13653116|r1909637 atatatggcactggggggcagacgtacacttc
                      ttataagcgactacgccactcggccgtatttta
                      gggcataggttggttttactccgccggtgccg

gi|3599509|gb|A      33                      GKPLMTQQQMSALSREGMLDALFAL
                      GKPP MTQQQMS LSREG+LDALF L
                      GKPPFMTQQQMSPLSREGILDALFVL
gi|13653116|r1909733 GTAACAG Intron 1 TAGgacctaaccctcctcgatggctgc
                      0-----[1909733:1916-0>gaccttcaaatacctcgagttactttt
                      gaactgtagggtttcaagaatccttc

gi|3599509|gb|A      59 FEECSPALMCKMKHVSSFVQK                      SD
                      FEECSPALMK+KHVS+FV+K                      SD
                      FEECSPALMKIKHVSNFVRK                      Y:Y[tat] SD
gi|13653116|r1916682 tggtagcgcaaaacgaatgcatGTAAGTT Intron 2 CAGATtg
                      taaggaccttataatgattga 1-----[1916746:1928-1> ca
                      taactgttggtgctgcctcgg                      cc

gi|3599509|gb|A      83 TIAELRELQPSARDFEVRSLVGCGHFAEVQVVREKATGDVYAMKIMKKK
                      TIAEL+ELQPSA+DFEVRSLVGCGHFAEVQVVREKATGD+YAMK+MKKK
                      TIAELQELQPSAKDFEVRSLVGCGHFAEVQVVREKATGDIYAMKVMKKK
gi|13653116|r1928115 aaggtcgccctgagtggaacggtgctggcgagagaggatgaagaaaa
                      ctcataataaccaatatggttgggatcatattgaaccgatactattaaa
                      catgaggcgtgagccacattatttcttagggaagaacgccttgaggggg

gi|3599509|gb|A      132 ALLAQEQ                      VSFFEEERNILSRSTSPWI
                      ALLAQEQ                      VSFFEEERNILSRSTSPWI
                      ALLAQEQ                      VSFFEEERNILSRSTSPWI

```

- 55/84 -

Fig. 20 (continued)

```

gi|13653116|r1928262 gttgcgcGTAGGAG Intron 3 TAGgtttgggcaattcaaacta
cttcaaa0-----[1928283:1935-0>tcttaaagattcggcgcggt
tagcggg tattgaggcaatacacggc

gi|3599509|gb|A 158 PQLQYAFQDKNNLYL VMEYQPGGDFL
PQLQYAFQDKN+LYL VMEYQPGGD+L
PQLQYAFQDKNHLYL VMEYQPGGDL
gi|13653116|r1935587 cctctgtcgaacctcGTGAGTC Intron 4 CAGgagtcggggtc
cataactaaaaatat0-----[1935632:1951-0>ttaaacggatt
caagtctgcatcttg cgatgtagcgg

gi|3599509|gb|A 184 SLLNRYEDQLDESMIQFYLAELILAVHSVHQMgyVH
SLLNRYEDQLDE++IQFYLAELILAVHSVH MGYVH
SLLNRYEDQLDENLIQFYLAELILAVHSVHLMGYVH
gi|13653116|r1951610 tctaattggctggacacttcggcatggcagccagtgc
cttagaaaataaattatatcatttctagtattgata
atgtatgcgatacagatcatggtgttccttggacgt

gi|3599509|gb|A 220 DIKPENILIDRTGEIKLVDFGSA
DIKPENIL+DRTG IKLVDFGSA
R:R[cga] DIKPENILVDRTGHIKLVDFGSA
gi|13653116|r1951718 CGGTAAGTG Intron 5 CAGAgaacgaacggcagcaacgggtgtg
2-----[1951720:1952-2> atacaatttagcgatattatgcc
ccgtgctctccaaccgggttatc

gi|3599509|gb|A 244 AKMNSNKV -DAKLPIGTPDYMAPEVL
AKMNSNK+ AKLPIGTPDYMAPEVL
AKMNSNKM VNAKLPIGTPDYMAPEVL
gi|13653116|r1953011 gaaataaaGTAAAAA Intron 6 TAGgagaccagacgtagcggc
catacaat0-----[1953035:1960-0>tacatctgccaatccatt
gagtacgg gtcacgtgcatcgtagg

```

- 56/84 -

Fig. 20 (continued)

gi|3599509|gb|A 269 TVMNEDRRGTYGLDCDWWSVGVVAYEMVYGKTPFTEGTSARTFNNIMNF
 TVMN D +GTYGLDCDWWSVGV+AYEM+YG++PF EGTSARTFNNIMNF
 TVMNGDGKGTYGLDCDWWSVGVIAIYEMIYGRSPFAEGTSARTFNNIMNF
 gi|13653116|r1960491 agaagggagatgcgtgtttgggagtgaatgatctgggatgaataaaaat
 cttagagagcagtagaggctgttcaattaggcctcagcccgcataattat
 tggcgtaaccccgcctcggagcgtctggttgaccagactcacctctgtc

gi|3599509|gb|A 318 Q RFLKFPDDPKVSSELLDLLQSLLCV
 Q RFLKFPDDPKVSS+ LDL+QSLLC
 Q RFLKFPDDPKVSSDFLDLIQSLLCG
 gi|13653116|r1960638 cGTAAAGA Intron 7 CAGcttatcggcagaagtcgcacatttg
 a0-----[1960641:1962-0>gttatcaacatggattattagttgg
 g gtgatatccagctctttgtacggcc

gi|3599509|gb|A 344 QKERLKFEGLCCHPFFARTDWNINRN
 QKERLKFEGLCCHPFF++ DWNINRN
 QKERLKFEGLCCHPFFSKIDWNINRN S:S[tct]
 gi|13653116|r1962909 cagacatggcttcctttaagtaaacaTGTAAGTA Intron 8
 aaagtatagtggaacttcatagaatga 1-----[1962988:19824
 gagaggtattccttcctatcgcttc

gi|3599509|gb|A 370 PPPFVPTLKSDDDTSNFDEPEKNSWAFILCVPAEPLAFSGEELP
 PPPFVPTLKSDDDTSNFDEPEKNSW P FSGEELP
 PPPFVPTLKSDDDTSNFDEPEKNSWSSSPCQLSPSGFSGEELP
 gi|13653116|r1982415 CAGCTccctgcacatgggatatggcgaattgtttctccactgttgggccc
 -1> cccttcctacaaaccataacaaacgtccccgatgccgtcgaatc
 tccctcccgtctccctttaaggtggtactgcggccaccgtaagg

gi|3599509|gb|A 415 FVGFSYSKALGYLGRS SVVSSLD
 FVGFSYSKALG LGRS SVVS LD
 FVGFSYSKALGILGRS E:E[gag] SVVSGLD
 gi|13653116|r1982552 tggtttaagcgacgatGAGTAAGTG Intron 9 TAGGtggtgcg
 ttgtcagactgttggc 2-----[1982602:2000-2> ctctcgta
 tggtgccgaggtttat ttggtgc

- 57/84 -

Fig. 20 (continued)

gi|3599509|gb|A 439 SPAKVSSMEKKLLIKSKELQDSQDKCHK
 SPAK SSMEKKLLIKSKELQDSQDKCHK
 SPAKTSSMEKKLLIKSKELQDSQDKCHK
 gi|13653116|r2000764 tcgaaatagaacaaaaagccgtcgatcaGTATTTA Intron 10
 cccacgctaaatttagaataacaaagaa0----- [2000848:20017
 ctcgtccgagatccacagaactgcgtcg

gi|3599509|gb|A 467 MEQEMTRLHRRVSEVEAVLSQKEVELKASETQRSLLLEQDLATYITE
 MEQEMTRLHRRVSEVEAVLSQKEVELKASETQRSLLLEQDLATYITE
 MEQEMTRLHRRVSEVEAVLSQKEVELKASETQRSLLLEQDLATYITE
 gi|13653116|r2001753 CAGagcgaactccagtgggggcacagggcagtgacatccgcgcgataag
 -0>taaatcgtaggtcatacttgaaatataccacagcttaaatccatca
 gggagcgatgagaggggtgttgggggggctgtgaccgggcttcccaa

gi|3599509|gb|A 513 CS SLKRSLEQARMEVSQEDDKALQLL
 CS SLKRSLEQARMEVSQEDDKALQLL
 CS SLKRSLEQARMEVSQEDDKALQLL
 gi|13653116|r2001894 taGTGAGCC Intron 11 CAGatacatgcgcaggtcgggagcccc
 gg0----- [2001900:2003-0>gtaggtaacgtatcaaaaactatt
 ct cagatggaaggggcggtcaaggctc

gi|3599509|gb|A 539 HDIREQSRKLQEIKEQ EYQAQVEEMR
 HDIREQSRKLQEIKEQ EYQAQVEEMR
 HDIREQSRKLQEIKEQ EYQAQVEEMR
 gi|13653116|r2003242 cgaagcacaccgaagcGTAGGCC Intron 12 TAGgtcgcgggaa
 aatgaaggataataaaa0----- [2003290:2008-0>aaacataatg
 ttcaggcggcaacagg gcgtagaagg

gi|3599509|gb|A 565 LMMNQLEEDLVSARRRSDLYESELRESRLAAEEFKRKANECQHKLK
 LMMNQLEEDLVSARRRSDLYESELRESRLAAEEFKRKA ECQHKL+K
 LMMNQLEEDLVSARRRSDLYESELRESRLAAEEFKRKATECQHKLK
 gi|13653116|r2008995 taaactgggcgtgaacagctgtgcagtcgggggtacagagtccacta
 ttttaataaattccgggggataacatgacgtccaatagaccagaaatta
 ggggtggagttcaaaagtccatggagtggttaacggagaatgtaggg

- 58/84 -

Fig. 20 (continued)

gi|3599509|gb|A 612

AKDQ GKPEVGEYSKLEK

AKDQ GKPEVGEY+KLEK

AKDQ GKPEVGEYAKLEK

gi|13653116|r2009136 GTAGTCA Intron 13 CAGgagcgacggggtgacga

0-----[2009136:2009-0>caaagacatgaacataa

tgtaggtagaatgaggg

gi|3599509|gb|A 629

INAEQQLKIQELQEKLEK

INAEQQLKIQELQEKLEK

INAEQQLKIQELQEKLEK

gi|13653116|r2009450 GTATACT Intron 14 TAGaaggcccaacgccgacga

0-----[2009450:2009-0>tacaaatataataataa

cttgggcatggcagaggg

gi|3599509|gb|A 647

AVKASTEATELLQNIROAKERAEREL

AVKASTEATELLQNIROAKERAEREL

AVKASTEATELLQNIROAKERAEREL

gi|13653116|r2010022 GTAAGCC Intron 15 TAGggagaaggagcccaaccgagcggagc

0-----[2010022:2012-0>ctacgcaccattaatgacaagcagat

taaccggccggggtccgaggacgggg

gi|3599509|gb|A 673 EKLHNREDSSEGIKKLVEAE

ERRHS

EKL NREDSSEGI+KKLVEAE

ERRHS

EKLQNRDSSEGIKKLVEAE

ERRHS

gi|13653116|r2012975 gaccacggttgaaaacggggGTGAGCA Intron 16 CAGgccct

aataagaaccagtgaattaca0-----[2013038:2014-0>aggac

ggggcagtttaccaggggatg acctt

gi|3599509|gb|A 699 LENKVKRLETMERRENRLKDDIQTKEQIQQMADKIL

LENKVKRLETMERRENRLKDDIQTKE+QIQQMADKIL

LENKVKRLETMERRENRLKDDIQTKEQIQQMADKIL

gi|13653116|r2014912 cgaagaacgaagcagaacaggacaatccaccaggaac

taaatagtactaggaagtaataacacaataatcaatt

ggcgagaagcggttaacagggtccgaacagcggttatg

- 59/84 -

Fig. 20 (continued)

gi|3599509|gb|A 736

ELEEKHREAQVSAQHLEVHLKQKEQH

ELEEKHREAQVSAQHLEVHLKQKEQH

ELEEKHREAQVSAQHLEVHLKQKEQH

gi|13653116|r2015023 GTGAGCA Intron 17 TAGgcgaccggcggtgcccgccacagcc

0-----[2015023:2018-0>ataaaagacatccaatatataaaaaa

gcagatggcacacgcaagcgagaggg

gi|3599509|gb|A 762 YEEKIK

VLDNQIKKDLADKESLENMM

YEEKIK

VLDNQIKKDLADKE+LENMM

YEEKIK

VLDNQIKKDLADKETLENMM

gi|13653116|r2018703 tggaaaGTAAAGA Intron 18 TAGgtgacaaagcggagacgaaa

aaaata0-----[2018721:2024-0>ttaaataaatcaaactaatt

tgagta

ggctgagacgtcggagggcg

gi|3599509|gb|A 788 QRHEEEAHEKGKILSEQKA

MINAMDS

QRHEEEAHEKGKILSEQKA

MINAMDS

QRHEEEAHEKGKILSEQKA

MINAMDS

gi|13653116|r2024812 cacggggcgagaaacagcagGTAGGTA Intron 19 CAGaaagagt

agaaaacaaagattgaaac0-----[2024869:2027-0>ttactac

gacgggctggcatccaggg

gcttgto

gi|3599509|gb|A 814 KIRSLEQRIVELSEANKLAANSSLFTQRN

KIRSLEQRIVELSEANKLAANSSLFTQRN

KIRSLEQRIVELSEANKLAANSSLFTQRN

gi|13653116|r2027128 aaatcgcaaggctggaacggaaactacaa

atgctaagttatcacatccaggttcaga

gcacgaggtgagtactataatctttcagc

gi|3599509|gb|A 843

KAQEEMISELRQQKFYLETQAGK

KAQEEMISELRQQKFYLETQAGK

M:M[atg]

KAQEEMISELRQQKFYLETQAGK

gi|13653116|r2027215 ATGTAAGTA Intron 20 CAGGagcggaatgcaccattcgacgga

2-----[2027217:2028-2> acaaattcatgaaatatacacga

gcaaggttacgagatcggagtg

- 60/84 -

Fig. 20 (continued)

gi|3599509|gb|A 867 LEAQNRLKEEQLEKISHQDHSKSRLLLELETRLRE
LEAQNRLKEEQLEKISHQDHSK+RLLLELETRLRE
LEAQNRLKEEQLEKISHQDHSKRNRLLELETRLRE
gi|13653116|r2028332 tggcacacggccgaaaccgcagaaccgcgaatcg
tacaagataaataatgaaaagaaagttatactga
ggcgcaagggggggccacctcgtgggaggaaggg

gi|3599509|gb|A 902 VSLEHHEQKLELKRQLTELQLSLQER
VSLEHHEQKLELKRQLTELQLSLQER
VSLEHHEQKLELKRQLTELQLSLQER
gi|13653116|r2028437 GTGAGAG Intron 21 CAGgacgcggcacgcacccagccctccgc
0-----[2028437:2033-0>tgtaaaaaatatagatcatatctaag
ctagcgggagggcgcgagagccgggc

gi|3599509|gb|A 928 ESQLTALQAARAALSQLRQAKTELETTAEAEIIIQALT
ESQLTALQAARAALSQLRQAKTELETTAEAEIIIQALT
ESQLTALQAARAALSQLRQAKTELETTAEAEIIIQALT
gi|13653116|r2033637 gtctagccggcgcgaccccgaaagcggaaggggggacgca
acatcctaccgcctagatgacacataaccacaaaatactc
gaggacggtaggcggcgctcgggaggagcaaataggcgacg

gi|3599509|gb|A 968 AHRDEIQRKFDALRNSCT
AHRDEIQRKFDALRNSCT
AHRDEIQRKFDALRNSCT
gi|13653116|r2033757 GTAGGTC Intron 22 TAGgcaggaccatggccaata
0-----[2033757:2043-0>cagaatagatactgaggc
atatacgcatTTTTCTT

gi|3599509|gb|A 986 VITDLEEQLNQLTEDNAELNNQNFYL
VITDLEEQLNQLTEDNAELNNQNFYL
VITDLEEQLNQLTEDNAELNNQNFYL
gi|13653116|r2043396 GTGAGTA Intron 23 TAGgaagcggccaccaggaggcaacattt
0-----[2043396:2050-0>ttcataaataatcaaacataaaatat
acacggggacggcgccctaccacccg

- 61/84 -

Fig. 20 (continued)

```

gi|3599509|gb|A 1012 SKQLDEASGANDEIVQLRSEVDHLRREITEREMQLTSQKQ
                      SKQLDEASGANDEIVQLRSEVDHLRREITEREMQLTSQKQ
                      SKQLDEASGANDEIVQLRSEVDHLRREITEREMQLTSQKQ
gi|13653116|r2050527 taccgggtggaggagcccagggccccgaagcgaccaaacac
                      caataaccgcaaattatggataatggatcagatatcgaaa
                      caactgttccccgtaagatagctccggcgaagggtccgga

gi|3599509|gb|A 1052                      TMEALKTTCTMLEEQVLDLEALNDEL
                      TMEALKTTCTMLEEQV+DLEALNDEL
                      TMEALKTTCTMLEEQVMDLEALNDEL
gi|13653116|r2050647 GTAAGGA Intron 24 CAGaaggcaaataacggcgagtggcaggc
                      0-----[2050647:2051-0>ctactaccgcttaaattataactaaat
                      ggggtggcgccgggagcgtggcactgg

gi|3599509|gb|A 1078 LEKERQWEAWRSVLGDEKSQFECRVRELQRMMLDTEKQS
                      LEKERQWEAWRSVLGDEKSQFECRVRELQRMMLDTEKQS
                      LEKERQWEAWRSVLGDEKSQFECRVRELQRMMLDTEKQS
gi|13653116|r2051527 cgagcctggtaagcgggatctgtcgcgccaacgagaca
                      taaagagacgggttgaaacataggtgatagttacaaag
                      aaagggggcgccgcttgacgtgtgtagggaggccgagc

gi|3599509|gb|A 1116                      ARADQRITESRQVVELAVKEHKA
                      ARADQRITESRQVVELAVKEHKA
                      R:R[agg] ARADQRITESRQVVELAVKEHKA
gi|13653116|r2051641 AGGTGGGGC Intron 25 CAGGgaggccaagtccgggaggagcag
                      2-----[2051643:2055-2> cgcaagtcacgattatctaaaac
                      gactggccgtcgggggagggcggt

gi|3599509|gb|A 1140 BILALQQALKEQKLKAESLSDK                      LNDL
                      BILALQQALKEQKLKAESLSDK                      LNDL
                      BILALQQALKEQKLKAESLSDK                      LNDL
gi|13653116|r2055246 gacgcccgcagcacaggactgaGTCAGCG Intron 26 TAGcagc
                      attctaactaaaatacagtcaa0-----[2055312:2057-0>taat
                      gtctgggtcagggggcgccctcg                      ctcg

```

- 62/84 -

Fig. 20 (continued)

gi|3599509|gb|A 1166 EKKHAMLEMNARSLQQKLETERELKQRLLEE
EKKHAMLEMNARSLQQKLETERELKQRLLEE
EKKHAMLEMNARSLQQKLETERELKQRLLEE
gi|13653116|r2057212 gaacgacgaagcatccacgagcgacaccgg
aaaacttatacggtaaatacagataagttaa
gggttgtagtcacaggggtaagcaggtgag

gi|3599509|gb|A 1197 QAKLQQQMDLQKNHIFRLTQGLQEAL
QAKLQQQMDLQKNHIFRLTQGLQEAL
QAKLQQQMDLQKNHIFRLTQGLQEAL
gi|13653116|r2057305 GTGAGTG Intron 27 TAGcgatcccagccaacatccacgccggc
0-----[2057305:2064-0>acataaatataaaaattgtcagtaact
acaaggggcgatctctgtaagaata

gi|3599509|gb|A 1223 DRADLLKTERS DLEYQLENIQ VLYSH
DRADLLKTERS DLEYQLENIQ VLYSH
DRADLLKTERS DLEYQLENIQ VLYSH
gi|13653116|r2064435 gcggccaagaagtgtccgaacGTGAGGA Intron 28 TAGgcttc
agcattacaggataaataa0-----[2064498:2065-0>ttaca
tgtaggaatcggtggactg tcttt

gi|3599509|gb|A 1249 EKVKMEGTISQQTKLIDFLQAKMDQPAK KKK
EKVKMEGTISQQTKLIDFLQAKMDQPAK KKK
EKVKMEGTISQQTKLIDFLQAKMDQPAK KKK
gi|13653116|r2065236 gagaaggaatccaacagtccgaagccgaaaa
aatatagctcaacattattacataacccaaaa
aggagactttaacactttgacagcattagag

gi|3599509|gb|A 1280 VPLQYNELKLALALEKEKARCAELEEEAL
VPLQYNELKLALALEKEKARCAELEEEAL
VPLQYNELKLALALEKEKARCAELEEEAL
gi|13653116|r2065329 GTGAGTC Intron 29 CAGgccctagcacgagagctggcgggc
0-----[2065329:2066-0>tctaaaatatctaaaacggcataact
ttggctggggcggggatctagagact

- 63/84 -

Fig. 20 (continued)

gi|3599509|gb|A 1306 QKTRIELRSAREE AHRKATDHPH
 QKTRIELRSAREE AHRKATDHPH
 QKTRIELRSAREE A:A[gct] AHRKATDHPH
 gi|13653116|r2067068 caacagcctgcggGGTAGGGG Intron 30 CAGCTgccagagccc
 aacgtatgccgaa 1-----[2067108:2067-1> cagaccaaca
 ggcccgcgcccga cccaagccac

gi|3599509|gb|A 1330 PSTPATARQQIAMSIVRSPEHQPSAMSLAPPSSRRKESSTPE
 PSTPATARQQIAMSIVRSPEHQPSAMSLAPPSSRRKESSTPE
 PSTPATARQQIAMSIVRSPEHQPSAMSLAPPSSRRKESSTPE
 gi|13653116|r2067429 ctacgagaccagatgagctcgcccagaaccgcctacaagttacg
 cccccccaatctccttgccaaacgctgttccccgggaacccca
 acgaccggggccgcccgggagcgctcgggcgacccaggtatag

gi|3599509|gb|A 1374 FSRRLKERMHHNIPHRFNVGLNM
 FSRRLKERMHHNIPHRFNVGLNM
 E:E[gaa] FSRRLKERMHHNIPHRFNVGLNM
 gi|13653116|r2067561 GGTACGTT Intron 31 CAGAAataccagcaccaaccctaggcaa
 1-----[2067562:2071-1> tgggtaagtaaatacgtatgtat
 ttgttgacgcctttcaccaagcg

gi|3599509|gb|A 1398 RATKCAVCLDTVHFGRQASKCL C
 RATKCAVCLDTVHFGRQASKCL C
 RATKCAVCLDTVHFGRQASKCL E:E[gaa] C
 gi|13653116|r2071543 cgaatggctcgagctgccgtatcGGTAAGAT Intron 32 TAGAAt
 gccagctgtactatggaccagt 1-----[2071610:2072-1> g
 acagttgtgtcgctacgacatc t

gi|3599509|gb|A 1422 QVMCHPKCSTCLPATCGLPAEYATHFTEAFCDKMNNSPGLQSKPEGSSSL
 QVMCHPKCSTCLPATCGLPAEYATHFTEAFCDKMNNSPGLQ+KEP SSL
 QVMCHPKCSTCLPATCGLPAEYATHFTEAFCDKMNNSPGLQTKEPSSSL
 gi|13653116|r2072163 cgatccattattcgatgtcggtgactaggttcgaaatcgccaagcaaat
 attgacagccgtcccgggtccaaccatcactggaataccgtacaacgggt
 gggtccgcccgcgaccccggttatcaccgcccctcagccatcgcgggccccg

- 64/84 -

Fig. 20 (continued)

gi|3599509|gb|A 1471 HLEGWMKVP NNKRGQQGWDRKYI
 HLEGWMKVP NNKRGQQGWDRKYI
 HLEGWMKVP R:R[agg] NNKRGQQGWDRKYI
 gi|13653116|r2072310 ccggttaagcAGGTACCAT Intron 33 CAGGaaacgccgtgaata
 ataggtatc 2-----[2072339:2072-2> aaaggaaggagaat
 cgaggggggc tcaaagacgcggct

gi|3599509|gb|A 1495 VLEGSKVLIYDNEARE GQRPVEE
 VLEGSKVLIYDNEARE GQRPVEE
 VLEGSKVLIYDNEARE A:A[gct] GQRPVEE
 gi|13653116|r2072555 gcggtagcatgaggagGGTAAATT Intron 34 AAGCTgcacggg
 ttagcattttaaaca 1-----[2072604:2073-1> gagctaa
 cggaacacctctacaa aggggaa

gi|3599509|gb|A 1519 FELCLPDGDVSIHGAVGASELANTAKA
 FELCLPDGDVSIHGAVGASELANTAKA
 FELCLPDGDVSIHGAVGASELANTAKA D:D[gat]
 gi|13653116|r2073134 tgctccggggtacgggggtgcgaagagGGTGAGGA Intron 35
 tatgtcagatctagctgccatcaccac 1-----[2073216:20734
 tggctccgtattttctttcacatacaa

gi|3599509|gb|A 1546 VPYILKMESHPTTCWPGRTRYLLAPSFPDKQRWVTALESVVAG
 VPYILKMESHPTTCWPGRTRYLLAPSFPDKQRWVTALESVVAG
 VPYILKMESHPTTCWPGRTRYLLAPSFPDKQRWVTALESVVAG
 gi|13653116|r2073456 TAGATgctacaagtcccaattcgaacttcgcatcgacctgagtgtgggg
 -1> tcattatacacaccggcggtattccgtcaaaggtcctacttcg
 cacagggatcgcccgcgacccgatccctcagcgcccaaatcat

gi|3599509|gb|A 1591 GRVSREKAEADA KLLGNSLLKLEGDD
 GRVSREKAEADA KLLGNSLLKLEGDD
 GRVSREKAEADA KLLGNSLLKLEGDD
 gi|13653116|r2073593 gagtagagggggGTGAGTA Intron 36 AAGaccgatccacggg
 ggtcgaacacac0-----[2073629:2075-0>attgacttatagaa
 gattgaaaattt agtaccggagattc

- 65/84 -

Fig. 20 (continued)

gi|3599509|gb|A 1617 RLD MNCTLPFSDQ VVLVGTEEGLYAL
 RLD MNCTLPFSDQ VVLVGTEEGLYAL
 RLD MNCTLPFSDQ VVLVGTEEGLYAL
 gi|13653116|r2075228 ccgaatacctagcGTAATGC Intron 37 CAGggtggaggggctgc
 gtatagctctgaa0-----[2075267:2075-0>ttttgcaagtact
 tacgccggcctcg ggggccgagcccg

gi|3599509|gb|A 1643 NVLKNSLTHIPGIGAVFQIYI IKDLEKLLMIA
 NVLKNSLTH+PGIGAVFQIYI IKDLEKLLMIA
 NVLKNSLTHVPGIGAVFQIYI IKDLEKLLMIA
 gi|13653116|r2075459 agtaatcacgcgagggtcataaagcgaccaag
 attaactcatcgctgttatattaataattttc
 tcgaccactcaataaccatttcgcgggacgaa

gi|3599509|gb|A 1675 EERALCLVDVKKVKQSLAQSHLP
 EERALCLVDVKKVKQSLAQSHLP
 G:G[gga] EERALCLVDVKKVKQSLAQSHLP
 gi|13653116|r2075555 GGTGTGAG Intron 38 CAGGAggcgctcggaagactcgctccc
 1-----[2075556:2077-1> aagctgttataataactcacatc
 aggagttgcggagagcgcgccgt

gi|3599509|gb|A 1699 AQP DVSPNIFEAVKGCHLFAAGK IEN
 AQP DVSPNIFEAVKGCHLF AGK IEN
 AQP DISPNIFEAVKGCHLFGAGK IEN
 gi|13653116|r2077559 gccgatcaatgggagtccttggaGTAAGCT Intron 39 CAGaga
 cacatccattactaggattgcga0-----[2077628:2081-0>taa
 cgccacacttatcgcccgtagcg tgc

gi|3599509|gb|A 1725 SLCICAAMPSKVVILRYNDNLSKYCIRK
 LCICAAMPSKVVILRYN+NLSKYCIRK
 GLCICAAMPSKVVILRYNENLSKYCIRK
 gi|13653116|r2081360 gctatggacaaggacctagacaattacaGTAAGTC Intron 40
 gtgtgcctcgattttgaaaatgaagtga0-----[2081444:20838
 gccctacgccacctccccacccacccga

- 66/84 -

Fig. 20 (continued)

```

gi|3599509|gb|A 1753      EIETSEPCSCIHFTNYSILIGTNKFYEIDMKQYTL
                        EIETSEPCSCIHFTNYSILIGTNKFYEIDMKQYTL+
                        EIETSEPCSCIHFTNYSILIGTNKFYEIDMKQYTLE
gi|13653116|r2083845 CAGgagatgctataactaataacagaaattgagaactacg
                        -0>ataccacgggtatcaagtttgcaataatataaaacta
                        gagcagccctcccctctcctactaccaccgggcgcg

gi|3599509|gb|A 1789      FLDKNDHSLAPAVFASSNSFPV
                        FLDKNDHSLAPAVFA+SSNSFPV
                        E:E[gaa]      FLDKNDHSLAPAVFAASSNSFPV
gi|13653116|r2083956 GGTAGGAC Intron 41 CAGAAtcgaagcttgcggtggttaatcg
                        1-----[2083957:2084-1> ttaaaaactcccttccccagctct
                        cgtgtctcgattgtcctccctc

gi|3599509|gb|A 1813 SIVQANSAGQREEYLLCFH      FGVF
                        SIVQ NSAGQREEYLLCFH      FGVF
                        SIVQVNSAGQREEYLLCFH      E:E[gaa]      FGVF
gi|13653116|r2084125 tagcgaaggccggttcttcGGTGAGTC Intron 42 CAGAAtggt
                        cttatagcgagaaattgta 1-----[2084183:2084-1> tggt
                        acgggccaggaggcggtcc      tagc

gi|3599509|gb|A 1837 VDSYGRRSRTDDLKWSRLPLAF      Y
                        VDSYGRRSRTDDLKWSRLPLAF      Y
                        VDSYGRRSRTDDLKWSRLPLAF      A:A[gcc]      Y
gi|13653116|r2084995 gggtgacacaggcatactctgtGGTACGTG Intron 43 CAGCct
                        tacagggggcaatagggctctct 1-----[2085062:2087-1> a
                        gttcaatccactcggtcatgct      c

gi|3599509|gb|A 1861 REPYLFVTHFNSLEVIEIQARSSL
                        REPYLFVTHFNSLEVIEIQARSS
                        REPYLFVTHFNSLEVIEIQARSSA      G:G[ggg]
gi|13653116|r2087757 agctctgactatcggagacgcttgGGGTAAGCA Intron 44
                        gacatttcatactattatacgccc 2-----[2087831:20879
                        aactgtgccccacaaatgcgaccaa

```

- 67/84 -

Fig. 20 (continued)

gi|3599509|gb|A 1885 SPARAYLEIPNPRYLGPASSGAIYLASSYQDKLRVICCKGNLVK
+PARAYL+IPNPRYLGPASSGAIYLASSYQDKLRVICCKGNLVK
TPARAYLDIPNPRYLGPASSGAIYLASSYQDKLRVICCKGNLVK
gi|13653116|r2087965 CAGGacgcgctcgacacctcgcgattggattgtttcgatagattagacga
-2> cccgcatatcacgatgcctccgctatcccaaatgttgagatta
ctcagcgccgcgcgctctcaagtcggcacgtaagctccgaccgg

gi|3599509|gb|A 1931 ESGTEQHRVPSTSR SPNKRGPPT
ESGTE HR PSTSR SPNKRGPPT
ESGTEHHRGPSTSR S:S[agc] SPNKRGPPT
gi|13653116|r2088104 gtgagcccgctatcAGGTAACCA Intron 45 CAGCacaacgccca
acgcaaaggccccc 2-----[2088148:2095-2> gcaaggccc
gcctaccgcgcccc cccgacacg

gi|3599509|gb|A 1955 YNEHITKRVASSPAPPEGPSHPREPSTPHRYRDREGRTelRRDKSPGRP
YNEHITKRVASSPAPPEGPSHPREPSTPHRYR EGRTelRRDKSPGRP
YNEHITKRVASSPAPPEGPSHPREPSTPHRYR--EGRTelRRDKSPGRP
gi|13653116|r2095380 tagcaaacggtacgcccgcacccgcaaccctc ggcagccagatcgcc
aaaatcagtcgccccagcgacgacgccagag aggcattggaaccggc
ccgcccgcgcccaggcacccccgagacaccccc gggcggcgcggttccc

gi|3599509|gb|A 2004 LEREKSPGRMLSTRRERSPGRLFEDSSRGRLPAGAVRTPLSQVNKQVWDQ
LEREKSPGRMLSTRRERSPGRLFEDSSRGRLPAGAVRTPLSQVNKV
LEREKSPGRMLSTRRERSPGRLFEDSSRGRLPAGAVRTPLSQVNKVRQH
gi|13653116|r2095521 cgcgatcgcaaacagctcgactggaaagcccggggaacctcgaagacc
tagaaccggttgcgagccggttaaggggtccgctgcctcataatgaa
ggaggcccgccggaggccgggtaccgcggtgacggcgccggcggt

gi|3599509|gb|A 2053 S
S
S
gi|13653116|r2095668 t
c
c

- 68/84 -

Fig. 20 (continued)

gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	match	1909637	2095670
3906.49				
+ . gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	1909637	1909732
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	1909733	1916603
0.00				
+ . gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	1916604	1916745
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	1916746	1928106
0.00				
+ . gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	1928107	1928282
0.00				
+ 2 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	1928283	1935529
0.00				
+ . gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	1935530	1935631
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	1935632	1951576
0.00				
+ . gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	1951577	1951719
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	1951720	1952940
0.00				
+ . gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	1952941	1953034
0.00				
+ 1 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	1953035	1960436
0.00				
+ . gi 3599509 gb AAC72823.1				

- 69/84 -

Fig. 20 (continued)

gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	1960437	1960640
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	1960641	1962833
0.00				
+ . gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	1962834	1962987
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	1962988	1982417
0.00				
+ . gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	1982418	1982601
0.00				
+ 2 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	1982602	2000741
0.00				
+ . gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2000742	2000847
0.00				
+ 1 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2000848	2001755
0.00				
+ . gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2001756	2001899
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2001900	2003169
0.00				
+ . gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2003170	2003289
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2003290	2008964
0.00				
+ . gi 3599509 gb AAC72823.1				

- 70/84 -

Fig. 20 (continued)

gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2008965	2009135
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2009136	2009398
0.00				
+ gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2009399	2009449
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2009450	2009967
0.00				
+ gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2009968	2010021
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2010022	2012896
0.00				
+ gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2012897	2013037
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2013038	2014896
0.00				
+ gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2014897	2015022
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2015023	2018624
0.00				
+ gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2018625	2018720
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2018721	2024751
0.00				
+ gi 3599509 gb AAC72823.1				

- 71/84 -

Fig. 20 (continued)

gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2024752	2024868
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2024869	2027106
0.00				
+ . gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2027107	2027216
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2027217	2028261
0.00				
+ . gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2028262	2028436
0.00				
+ 1 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2028437	2033558
0.00				
+ . gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2033559	2033756
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2033757	2043341
0.00				
+ . gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2043342	2043395
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2043396	2050448
0.00				
+ . gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2050449	2050646
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2050647	2051448
0.00				
+ . gi 3599509 gb AAC72823.1				

- 72/84 -

Fig. 20 (continued)

gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2051449	2051642
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2051643	2055175
0.00				
+ . gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2055176	2055311
0.00				
+ 1 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2055312	2057199
0.00				
+ . gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2057200	2057304
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2057305	2064356
0.00				
+ . gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2064357	2064497
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2064498	2065220
0.00				
+ . gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2065221	2065328
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2065329	2066989
0.00				
+ . gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2066990	2067107
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2067108	2067396
0.00				
+ . gi 3599509 gb AAC72823.1				

- 73/84 -

Fig. 20 (continued)

gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2067397	2067561
0.00				
+ 2 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2067562	2071471
0.00				
+ gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2071472	2071609
0.00				
+ 2 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2071610	2072157
0.00				
+ gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2072158	2072338
0.00				
+ 2 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2072339	2072511
0.00				
+ gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2072512	2072603
0.00				
+ 1 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2072604	2073110
0.00				
+ gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2073111	2073215
0.00				
+ 2 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2073216	2073458
0.00				
+ gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2073459	2073628
0.00				
+ 2 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2073629	2075185
0.00				
+ gi 3599509 gb AAC72823.1				

- 74/84 -

Fig. 20 (continued)

gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2075186	2075266
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2075267	2075419
0.00				
+ gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2075420	2075555
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2075556	2077487
0.00				
+ gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2077488	2077627
0.00				
+ 2 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2077628	2081350
0.00				
+ gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2081351	2081443
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2081444	2083847
0.00				
+ gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2083848	2083956
0.00				
+ 0 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2083957	2084053
0.00				
+ gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2084054	2084182
0.00				
+ 2 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2084183	2084980
0.00				
+ gi 3599509 gb AAC72823.1				

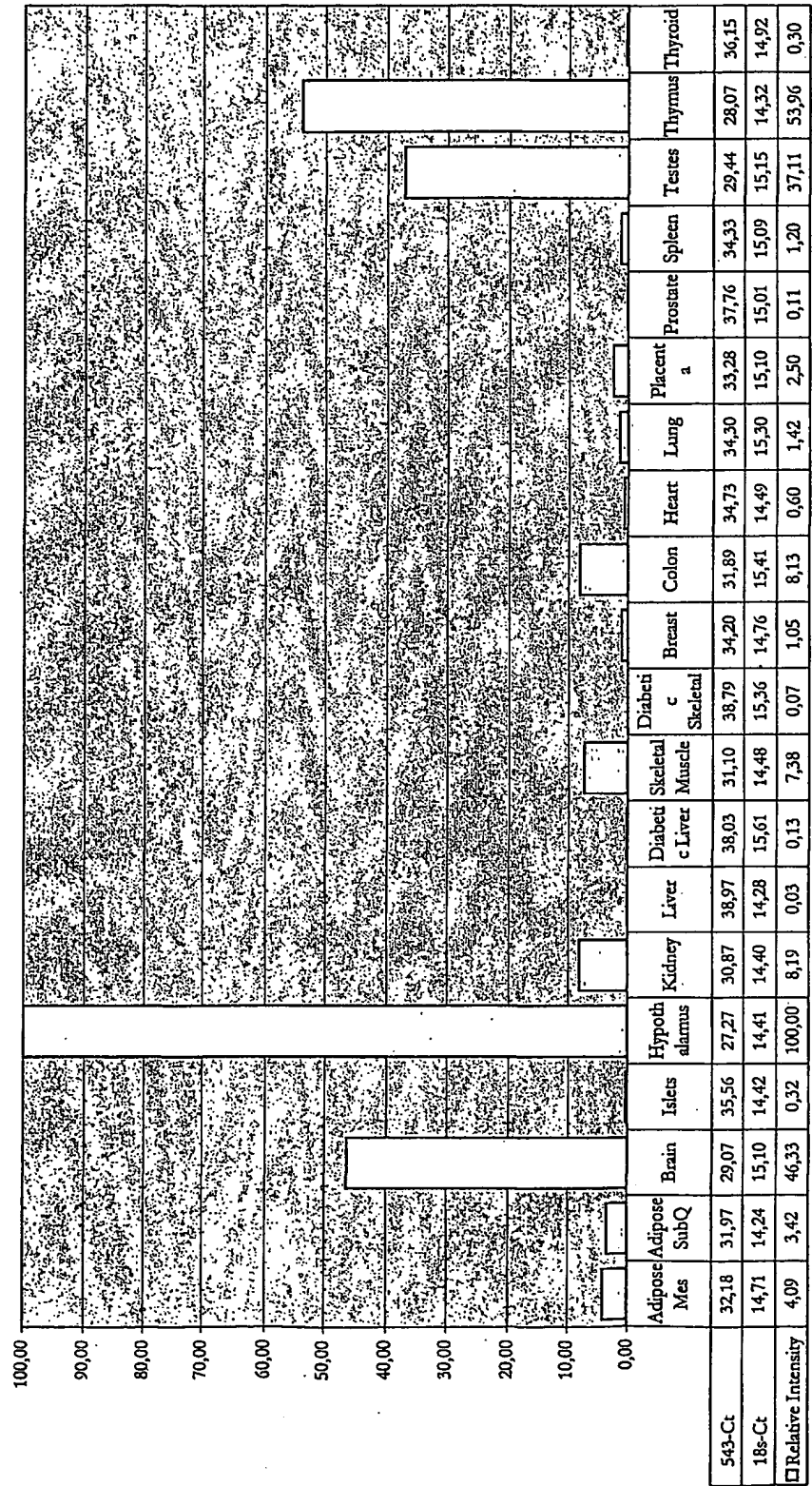
- 75/84 -

Fig. 20 (continued)

gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2084981	2085061
0.00				
+ 2 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2085062	2087751
0.00				
+ gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2087752	2087830
0.00				
+ 2 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2087831	2087967
0.00				
+ gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2087968	2088147
0.00				
+ 1 gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	intron	2088148	2095351
0.00				
+ gi 3599509 gb AAC72823.1				
gi 13653116 ref NT_009775.3 Hs12_9932	GeneWise	cds	2095352	2095670
0.00				
+ 1 gi 3599509 gb AAC72823.1				

FIG. 21

LBRI 543: Relative Expression



- 77/84 -

Fig. 22

TBLASTN - alignment of 543_Protein against emnew|AX166510|AX166510
Sequence 1 from Patent WO0138503.//:gbnew|AX166510|AX166510 Sequence 1 from
Patent WO0138503.

This hit is scoring at : 0.0 (expectation value)
Alignment length (overlap) : 2053
Identities : 99 %
Scoring matrix : BLOSUM62 (used to infer consensus pattern)
Hit reading frame : +1
Database searched : nrnee_1_;

Q: 1 MLKFKYGARNPLDAGAAEPIASRASRLNLFQGGKPPFMTQQQMSPLSREGILDALFVLFE
MLKFKYGARNPLDAGAAEPIASRASRLNLFQGGKPPFMTQQQMSPLSREGILDALFVLFE
H: 1 MLKFKYGARNPLDAGAAEPIASRASRLNLFQGGKPPFMTQQQMSPLSREGILDALFVLFE

ECSQPALMKIKHVSNFVRK-YSDTIAELQELQPSAKDFEVRSLVGCCHFAEVQVVREKAT
ECSQPALMKIKHVSNFV : YSDTIAELQELQPSAKDFEVRSLVGCCHFAEVQVVREKAT
ECSQPALMKIKHVSNFVPEVYSDTIAELQELQPSAKDFEVRSLVGCCHFAEVQVVREKAT

GDIYAMKVMKKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKNHLYLVMEYQPG
GDIYAMKVMKKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKNHLYLVMEYQPG
GDIYAMKVMKKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKNHLYLVMEYQPG

GDLLSLLNRYEDQLDENLIQFYLAELILAVHSVHLMGYVHRDIKPENILVDRTGHIKLVD
GDLLSLLNRYEDQLDENLIQFYLAELILAVHSVHLMGYVHRDIKPENILVDRTGHIKLVD
GDLLSLLNRYEDQLDENLIQFYLAELILAVHSVHLMGYVHRDIKPENILVDRTGHIKLVD

FGSAAKMNSNKMVNAKLPIGTPDYMAPEVLTVMNGDGKGTYGLDCDWWSVGVIAYEMIYG
FGSAAKMNSNKMVNAKLPIGTPDYMAPEVLTVMNGDGKGTYGLDCDWWSVGVIAYEMIYG
FGSAAKMNSNKMVNAKLPIGTPDYMAPEVLTVMNGDGKGTYGLDCDWWSVGVIAYEMIYG

RSPFAEGTSARTFNNIMNFQRFKFPDDPKVSSDFLDLIQSLLCGQKERLKFEGLCCHPF
RSPFAEGTSARTFNNIMNFQRFKFPDDPKVSSDFLDLIQSLLCGQKERLKFEGLCCHPF
RSPFAEGTSARTFNNIMNFQRFKFPDDPKVSSDFLDLIQSLLCGQKERLKFEGLCCHPF

FSKIDWNNIRNSPPPFVPTLKSDDDTSNFDEPEKNSWVSSSPCQLSPSGFSGEELPFVGF
FSKIDWNNIRNSPPPFVPTLKSDDDTSNFDEPEKNSWVSSSPCQLSPSGFSGEELPFVGF
FSKIDWNNIRNSPPPFVPTLKSDDDTSNFDEPEKNSWVSSSPCQLSPSGFSGEELPFVGF

- 78/84 -

Fig. 22 (continued)

SYSKALGILGRSESVVSGLDSPAKTSSMEKKLLIKSKELQDSQDKCHKMEQEMTRLHRRV
SYSKALGILGRSESVVSGLDSPAKTSSMEKKLLIKSKELQDSQDKCHKMEQEMTRLHRRV
SYSKALGILGRSESVVSGLDSPAKTSSMEKKLLIKSKELQDSQDKCHKMEQEMTRLHRRV

SEVEAVLSQKEVELKASETQRSLL EQDLATYITECSSLKRSLEQARMEVSQEDDKALQLL
SEVEAVLSQKEVELKASETQRSLL EQDLATYITECSSLKRSLEQARMEVSQEDDKALQLL
SEVEAVLSQKEVELKASETQRSLL EQDLATYITECSSLKRSLEQARMEVSQEDDKALQLL

HDIREQSRKLQEIKEQEYQAQVEEMRLMMNQLEEDLVSARRRSDLYESELRESRLAAEF
HDIREQSRKLQEIKEQEYQAQVEEMRLMMNQLEEDLVSARRRSDLYESELRESRLAAEF
HDIREQSRKLQEIKEQEYQAQVEEMRLMMNQLEEDLVSARRRSDLYESELRESRLAAEF

KRKATECQHKLKAKDQGKPEVG EYAKLEKINAEQQLKI QELQEKLEKAVKASTEATELL
KRKATECQHKLKAKDQGKPEVG EYAKLEKINAEQQLKI QELQEKLEKAVKASTEATELL
KRKATECQHKLKAKDQGKPEVG EYAKLEKINAEQQLKI QELQEKLEKAVKASTEATELL

QNIRQAKERAERELEKLQNREDSSEGIRKKLVEAEERRHSLENKVKRLETMERRENRLKD
QNIRQAKERAERELEKLQNREDSSEGIRKKLVEAEERRHSLENKVKRLETMERRENRLKD
QNIRQAKERAERELEKLQNREDSSEGIRKKLVEAEERRHSLENKVKRLETMERRENRLKD

DIQTKSQIQQMADKILELEEKHREAQVSAQHLEVHLKQKEQH YEEKIKVLDNQIKKDLA
DIQTKSQIQQMADKILELEEKHREAQVSAQHLEVHLKQKEQH YEEKIKVLDNQIKKDLA
DIQTKSQIQQMADKILELEEKHREAQVSAQHLEVHLKQKEQH YEEKIKVLDNQIKKDLA

DKETLENMMQRHEEEAHEKGKILSEQKAMINAMDSKIRSLEQRIVELSEANKLAANSSLF
DKETLENMMQRHEEEAHEKGKILSEQKAMINAMDSKIRSLEQRIVELSEANKLAANSSLF
DKETLENMMQRHEEEAHEKGKILSEQKAMINAMDSKIRSLEQRIVELSEANKLAANSSLF

TQRNMKAQEEMISELRQQKFYLETQAGKLEAQNRKLEE QLEKISHQDHS DKNRLLLELETR
TQRNMKAQEEMISELRQQKFYLETQAGKLEAQNRKLEE QLEKISHQDHS DKNRLLLELETR
TQRNMKAQEEMISELRQQKFYLETQAGKLEAQNRKLEE QLEKISHQDHS DKNRLLLELETR

LREVSLEHEEQKLELKRQLTELQLSLQERESQLTALQAARA ALESQLRQAKTELEETTAE
LREVSLEHEEQKLELKRQLTELQLSLQERESQLTALQAARA ALESQLRQAKTELEETTAE
LREVSLEHEEQKLELKRQLTELQLSLQERESQLTALQAARA ALESQLRQAKTELEETTAE

AEEBIQALTAHRDEIQRKFDALRNSCTVITDLEEQLNQLTEDNAELNNQNFYLSKQLDEA
AEEBIQALTAHRDEIQRKFDALRNSCTVITDLEEQLNQLTEDNAELNNQNFYLSKQLDEA
AEEBIQALTAHRDEIQRKFDALRNSCTVITDLEEQLNQLTEDNAELNNQNFYLSKQLDEA

- 79/84 -

Fig. 22 (continued)

SGANDEIVQLRSEVDHLRREITEREMQLTSQKQTMEALKTTCTMLEEQVMDLEALNDELL
SGANDEIVQLRSEVDHLRREITEREMQLTSQKQTMEALKTTCTMLEEQVMDLEALNDELL
SGANDEIVQLRSEVDHLRREITEREMQLTSQKQTMEALKTTCTMLEEQVMDLEALNDELL

EKERQWEAWRSVLGDEKSQFECRVRELQRM LDTEKQSRARADQRITESRQVVELAVKEHK
EKERQWEAWRSVLGDEKSQFECRVRELQRM LDTEKQSRARADQRITESRQVVELAVKEHK
EKERQWEAWRSVLGDEKSQFECRVRELQRM LDTEKQSRARADQRITESRQVVELAVKEHK

AEILALQQALKEQKLKAESLSDKLN DLEKKHAMLEMNARSLQQKLETERELKQRLLEEQA
AEILALQQALKEQKLKAESLSDKLN DLEKKHAMLEMNARSLQQKLETERELKQRLLEEQA
AEILALQQALKEQKLKAESLSDKLN DLEKKHAMLEMNARSLQQKLETERELKQRLLEEQA

KLQQQMDLQKNHIFRLTQGLQEALDRADLLKTERSDLEYQLENIQVLYSHEKVMEGTIS
KLQQQMDLQKNHIFRLTQGLQEALDRADLLKTERSDLEYQLENIQVLYSHEKVMEGTIS
KLQQQMDLQKNHIFRLTQGLQEALDRADLLKTERSDLEYQLENIQVLYSHEKVMEGTIS

QQTKLIDFLQAKMDQPAKKKKVPLQYNELKLALEKEKARCAELEELALQKTRIELRSAREE
QQTKLIDFLQAKMDQPAKKKKVPLQYNELKLALEKEKARCAELEELALQKTRIELRSAREE
QQTKLIDFLQAKMDQPAKKKKVPLQYNELKLALEKEKARCAELEELALQKTRIELRSAREE

AAHRKATDHPHPSTPATARQQIAMS AIVRSPEHQPSAMSLAPPSSRRKESSTPEEFSRR
AAHRKATDHPHPSTPATARQQIAMS AIVRSPEHQPSAMSLAPPSSRRKESSTPEEFSRR
AAHRKATDHPHPSTPATARQQIAMS AIVRSPEHQPSAMSLAPPSSRRKESSTPEEFSRR

LKERMHNNI PHRFNVGLNMRATKCAVCLDTVHFGRQASKCLECQVMCHPKCSTCLPATCG
LKERMHNNI PHRFNVGLNMRATKCAVCLDTVHFGRQASKCLECQVMCHPKCSTCLPATCG
LKERMHNNI PHRFNVGLNMRATKCAVCLDTVHFGRQASKCLECQVMCHPKCSTCLPATCG

LPAEYATHFTEAFCRDKMNSPGLQTKEPSSSLHLEGWMKVPRNNKRGQQGWRKYIVLEG
LPAEYATHFTEAFCRDKMNSPGLQTKEPSSSLHLEGWMKVPRNNKRGQQGWRKYIVLEG
LPAEYATHFTEAFCRDKMNSPGLQTKEPSSSLHLEGWMKVPRNNKRGQQGWRKYIVLEG

SKVLIYDNEAREAGQRPVEEFELCLPDGDVSIHGAVGASELANTAKADVPIYILKMESH PH
SKVLIYDNEAREAGQRPVEEFELCLPDGDVSIHGAVGASELANTAKADVPIYILKMESH PH
SKVLIYDNEAREAGQRPVEEFELCLPDGDVSIHGAVGASELANTAKADVPIYILKMESH PH

TTCWPGRTLYLLAPSFDPKQRWVTALESVVAGGRVSREKAEADAKLLGNSLLKLEGDDRL
TTCWPGRTLYLLAPSFDPKQRWVTALESVVAGGRVSREKAEADAKLLGNSLLKLEGDDRL
TTCWPGRTLYLLAPSFDPKQRWVTALESVVAGGRVSREKAEADAKLLGNSLLKLEGDDRL

- 80/84 -

Fig. 22 (continued)

DMNCTLPFSDQVVLVGTEEGLYALNVLKNSLTHVPGIGAVFQIYIIKDLEKLLMIAGEER
DMNCTLPFSDQVVLVGTEEGLYALNVLKNSLTHVPGIGAVFQIYIIKDLEKLLMIAGEER
DMNCTLPFSDQVVLVGTEEGLYALNVLKNSLTHVPGIGAVFQIYIIKDLEKLLMIAGEER

ALCLVDVKKVKQSLAQSHLPÄQPDISPNI FEAVKGCHLFGAGKIENGLCICAAMPSKVVI
ALCLVDVKKVKQSLAQSHLPÄQPDISPNI FEAVKGCHLFGAGKIENGLCICAAMPSKVVI
ALCLVDVKKVKQSLAQSHLPÄQPDISPNI FEAVKGCHLFGAGKIENGLCICAAMPSKVVI

LRYNENLSKYCIRKEIETSEPCSCIHFTNYSILIGTNKFYEIDMKQYTLEEFLDKNDHSL
LRYNENLSKYCIRKEIETSEPCSCIHFTNYSILIGTNKFYEIDMKQYTLEEFLDKNDHSL
LRYNENLSKYCIRKEIETSEPCSCIHFTNYSILIGTNKFYEIDMKQYTLEEFLDKNDHSL

APAVFAASSNSFPVSIVQVNSAGQREEYLLCFHEFGVFVDSYGRRSRTDDLKWSRLPLAF
APAVFAASSNSFPVSIVQVNSAGQREEYLLCFHEFGVFVDSYGRRSRTDDLKWSRLPLAF
APAVFAASSNSFPVSIVQVNSAGQREEYLLCFHEFGVFVDSYGRRSRTDDLKWSRLPLAF

AYREPYLFVTHFNSLEVIEIQARSSAGTPARAYLDIPNPRYLGPASSGAIYLASSYQDK
AYREPYLFVTHFNSLEVIEIQARSSAGTPARAYLDIPNPRYLGPASSGAIYLASSYQDK
AYREPYLFVTHFNSLEVIEIQARSSAGTPARAYLDIPNPRYLGPASSGAIYLASSYQDK

LRVICCKGNLVKESGTEHHRGPSTSRSSPNKRGPPPTYNEHITKRVASSPAPPEGPSHPRE
LRVICCKGNLVKESGTEHHRGPSTSRSSPNKRGPPPTYNEHITKRVASSPAPPEGPSHPRE
LRVICCKGNLVKESGTEHHRGPSTSRSSPNKRGPPPTYNEHITKRVASSPAPPEGPSHPRE

PSTPHRYREGRTTELRRDKSPGRPLEREKSPGRMLSTRRERSPGRLFEDSSRGRLPAGAVR
PSTPHRYREGRTTELRRDKSPGRPLEREKSPGRMLSTRRERSPGRLFEDSSRGRLPAGAVR
PSTPHRYREGRTTELRRDKSPGRPLEREKSPGRMLSTRRERSPGRLFEDSSRGRLPAGAVR

TPLSQVNKVDQS 2052

TPLSQVNKV . . S

TPLSQVNKVRQHS 6159

- 81/84 -

Fig. 23

TBLASTN - alignment of 543_Protein against BAYER_LIB_DNA|wu_373006001280181
Bayer Corp Pharma Proprietary OP Library: Fat Rat Hypothalamus Linda Wu Fr
Oct 15 15:45:51 EDT 1999

This hit is scoring at : 2e-37 (expectation value)

Alignment length (overlap) : 77

Identities : 100 %

Scoring matrix : BLOSUM62 (used to infer consensus pattern)

Hit reading frame : -3

Database searched : bayerall_1_;

Q: 964 IQALTAHRDEIQRKFDALRNSCTVITDLEEQLNQLTEDNAELNNQNFYLSKQLDEASGAN
IQALTAHRDEIQRKFDALRNSCTVITDLEEQLNQLTEDNAELNNQNFYLSKQLDEASGAN
H: 231 IQALTAHRDEIQRKFDALRNSCTVITDLEEQLNQLTEDNAELNNQNFYLSKQLDEASGAN

DEIVQLRSEVDHLRREI 1040

DEIVQLRSEVDHLRREI

DEIVQLRSEVDHLRREI 1

Fig. 24

ATGTTGAAGTTCAAATATGGAGCGCGGAATCCTTTGGATGCTGGTGCTGCTGAACCCATT
GCCAGCCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACCACCCCTTTATGACTCAA
CAGCAGATGTCTCTCTTTCCCGAGAAGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA
GAATGCAGTCAGCCTGCTCTGATGAAGATTAAGCACGTGAGCAACTTTGTCCGGAAGTAT
TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCAAAGGACTTCGAAGTCAGA
AGTCTTGTAGGTTGTGGTCACCTTTGCTGAAGTGCAGGTGGTAAGAGAGAAAGCAACCGGG
GACATCTATGCTATGAAAGTGATGAAGAAGAAGGCTTTATTGGCCCAGGAGCAGGTTTCA
TTTTTTGAGGAAGAGCGGAACATATTATCTCGAAGCACAAAGCCCGTGGATCCCCCAATTA
CAGTATGCCTTTTCAGGACAAAAATCACCTTTATCTGGTCATGGAATATCAGCCTGGAGGG
GACTTGCTGTCACTTTTGAATAGATATGAGGACCAGTTAGATGAAAACCTGATACAGTTT
TACCTAGCTGAGCTGATTTTGGCTGTTTACAGCGTTCATCTGATGGGATACGTGCATCGA
GACATCAAGCCTGAGAACAATTCTCGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT
GGATCTGCCGCGAAATGAATTCAAACAAGATGGTGAATGCCAAACTCCCGATTGGGACC
CCAGATTACATGGCTCCTGAAGTGCTGACTGTGATGAACGGGGATGGAAAAGGCACCTAC
GGCCTGGACTGTGACTGGTGGTCAGTGGGCGTGATTGCCTATGAGATGATTTATGGGAGA
TCCCCCTTCGAGAGGGGAACCTCTGCCAGAACCTTCAATAACATTATGAATTTCCAGCGG
TTTTTGAATTTCCAGATGACCCCAAAGTGAGCAGTGACTTTCTTGATCTGATTCAAAGC
TTGTTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGGTCTTTGCTGCCATCCTTTCTTC
TCTAAAATTGACTGGAACAACATTCGTAACCTCTCCTCCCCCTTCGTTCCACCCCTCAAG
TCTGACGATGACACCTCCAATTTTGATGAACCAGAGAAGAATTTCGTGGGTTTCATCCTCT
CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGTGGGGTTTTCG
TACAGCAAGGCACTGGGGATTCTTGGTAGATCTGAGTCTGTTGTGTCGGGTCTGGACTCC
CCTGCCAAGACTAGCTCCATGGAAAAGAACTTCTCATCAAAGCAAAGAGCTACAAGAC
TCTCAGGACAAGTGTCACAAGATGGAGCAGGAAATGACCCGGTTACATCGGAGAGTGTC
GAGGTGGAGGCTGTGCTTAGTCAGAAGGAGGTGGAGCTGAAGGCCTCTGAGACTCAGAGA
TCCCTCCTGGAGCAGGACCTTGCTACCTACATCACAGAATGCAGTAGCTTAAAGCGAAGT
TTGGAGCAAGCACGGATGGAGGTGTCCAGGAGGATGACAAAGCACTGCAGCTTCTCCAT
GATATCAGAGAGCAGAGCCGGAAGCTCCAAGAAATCAAAGAGCAGGAGTACCAGGCTCAA
GTGGAAGAAATGAGGTTGATGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAAGA
CGGAGTGATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGCTGCTGAAGAATTCAAG
CGGAAAGCGACAGAATGTGAGCATAAACTGTTGAAGGCTAAGGATCAAGGGAAGCCTGAA
GTGGGAGAATATGCGAAACTGGAGAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG
CTCCAAGAGAACTGGAGAAGGCTGTAAAAGCCAGCACGGAGGCCACCGAGCTGCTGCAG
AATATCCGCCAGGCAAAGGAGCGAGCCGAGAGGGAGCTGGAGAAGCTGCAGAACCAGAG
GATTCTTCTGAAGGCATCAGAAAGAAGCTGGTGAAGCTGAGGAACGCCGCCATTCTCTG
GAGAACAAGGTAAAGAGACTAGAGACCATGGAGCGTAGAGAAAACAGACTGAAGGATGAC
ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATCTGGAGCTCGAAGAG
AAACATCGGGAGGCCCAAGTCTCAGCCAGCACCTAGAAGTGCACCTGAAAACAGAAAGAG

Fig. 24 (continued)

CAGCACTATGAGGAAAAGATTAAAGTGTGGACAATCAGATAAAGAAAGACCTGGCTGAC
AAGGAGACACTGGAGAACATGATGCAGAGACACGAGGAGGAGGCCCATGAGAAGGGCAA
ATTCTCAGCGAACAGAAGGCGATGATCAATGCTATGGATTCCAAGATCAGATCCCTGGAA
CAGAGGATTGTGGAACGTCTGAAGCCAATAAACTTGCAGCAAATAGCAGTCTTTTTTACC
CAAAGGAACATGAAGGCCCAAGAAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC
CTGGAGACACAGGCTGGGAAGTTGGAGGCCCAGAACCAGAACTGGAGGAGCAGCTGGAG
AAGATCAGCCACCAAGACCACAGTGACAAGAATCGGCTGCTGGAACCTGGAGACAAGATTG
CGGGAGGTCAGTCTAGAGCACGAGGAGCAGAACTGGAGCTCAAGCGCCAGCTCACAGAG
CTACAGCTCTCCCTGCAGGAGCGCGAGTCACAGTTGACAGCCCTGCAGGCTGCACGGGGC
GCCCTGGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACCACAGCAGAAGCT
GAAGAGGAGATCCAGGCACTCACGGCACATAGAGATGAAATCCAGCGCAAATTTGATGCT
CTTCGTAACAGCTGTACTGTAATCACAGACCTGGAGGAGCAGCTAAACCAGCTGACCGAG
GACAACGCTGAACTCAACAACCAAACTTCTACTTGTCCAAACAACCTCGATGAGGCTTCT
GGCGCCAACGACGAGATTGTACAACCTGCGAAGTGAAGTGGACCATCTCCGCCGGGAGATC
ACGGAACGAGAGATGCAGCTTACCAGCCAGAAGCAAACGATGGAGGCTCTGAAGACCACG
TGCACCATGCTGGAGGAACAGGTCATGGATTGGAGGCCCTAAACGATGAGCTGCTAGAA
AAAGAGCGGCAGTGGGAGGCCTGGAGGAGCGTCTGGGTGATGAGAAATCCAGTTTGAG
TGTCGGGTTTCGAGAGCTGCAGAGAATGCTGGACACCGAGAAACAGAGCAGGGCGAGAGCC
GATCAGCGGATCACCGAGTCTCGCCAGGTGGTGGAGCTGGCAGTGAAGGAGCACAAGGCT
GAGATTCTCGCTCTGCAGCAGGCTCTCAAAGAGCAGAAGCTGAAGGCCGAGAGCCTCTCT
GACAAGCTCAATGACCTGGAGAAGAAGCATGCTATGCTTGAAATGAATGCCCGAAGCTTA
CAGCAGAAGCTGGAGACTGAACGAGAGCTCAAACAGAGGCTTCTGGAAGAGCAAGCCAAA
TTACAGCAGCAGATGGACCTGCAGAAAAATCACATTTTCCGTCTGACTCAAGGACTGCAA
GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAAGTGACTTGGAGTATCAGCTG
GAAAACATTGAGTTCTCTATTCTCATGAAAAGGTGAAATGGAAGGCACTATTTCTCAA
CAAACCAAACTCATTGATTTTCTGCAAGCCAAAATGGACCAACCTGCTAAAAAGAAAAAG
GTTCTCTGTCAGTACAATGAGCTGAAGCTGGCCCTGGAGAAGGAGAAAGCTCGCTGTGCA
GAGCTAGAGGAAGCCCTTCAGAAGACCCGCATCGAGCTCCGGTCCGCCCGGAGGAAGCT
GCCCACCGCAAAGCAACGGACCACCCACACCCATCCACGCCAGCCACCGCGAGGCAGCAG
ATCGCCATGTCCGCCATCGTGCGGTGCGCAGAGCACCAGCCAGTGCCATGAGCCTGCTG
GCCCCGCCATCCAGCCGAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGTCTT
AAGGAACGCATGCACCACAATATTCCTCACCGATTCAACGTAGGACTGAACATGCGAGCC
ACAAAGTGTGCTGTGTGCTGGATACCGTGCACTTTGGACGCCAGGCATCCAAATGTCTC
GAATGTCAGGTGATGTGTACCCCAAGTGCTCCACGTGCTTGCCAGCCACCTGCGGCTTG
CCTGCTGAATATGCCACACACTTCACCGAGGCCTTCTGCCGTGACAAAATGAACTCCCCA
GGTCTCCAGACCAAGGAGCCAGCAGCAGCTTGACCTGGAAGGGTGGATGAAGGTGCCC
AGGAATAACAAACGAGGACAGCAAGGCTGGGACAGGAAGTACATTGTCTGGAGGGATCA
AAAGTCCTCATTATGACAATGAAGCCAGAGAAGCTGGACAGAGGCCGGTGAAGAATTT

Fig. 24 (continued)

GAGCTGTGCCTTCCCCACGGGGATGTATCTATTTCATGGTGCCGTGGTGCTTCCGAATC
GCAAATACAGCCAAAGCAGATGTCCCATACATACTGAAGATGGAATCTCACCCGCACACC
ACCTGCTGGCCCGGGAGAACCCTCTACTTGCTAGCTCCCAGCTTCCCTGACAAACAGCGC
TGGGTCAACCGCCTTAGAATCAGTTGTTCGCAGGTGGGAGAGTTTCTAGGGAAAAAGCAGAA
GCTGATGCTAAACTGCTTGGAACCTCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGAC
ATGAACTGCACGCTGCCCTTCAGTGACCAGGTGGTGTGGTGGGCACCGAGGAAGGGCTC
TACGCCCTGAATGTCTTGAAAACTCCCTAACCCATGTCCCAGGAATTGGAGCAGTCTTC
CAAATTTATATTATCAAGGACCTGGAGAAGCTACTCATGATAGCAGGAGAAGAGCGGGCA
CTGTGTCTTGTGGACGTGAAGAAAGTGAACAGTCCCTGGCCCAGTCCCACCTGCCTGCC
CAGCCCAGCATCTCACCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTGTGTTGGGGCA
GGCAAGATTGAGAACGGGCTCTGCATCTGTGCAGCCATGCCAGCAAAGTCGTCAATCTC
CGCTACAACGAAAACCTCAGCAAATACTGCATCCGGAAAGAGATAGAGACCTCAGAGCCC
TGCAGCTGTATCCACTTCACCAATTACAGTATCCTCATTTGGAACCAATAAATTCTACGAA
ATCGACATGAAGCAGTACACGCTCGAGGAATTCCTGGATAAGAATGACCATTCTTGGCA
CCTGCTGTGTTTGCCGCCTCTTCCAACAGCTTCCCTGTCTCAATCGTGCAGGTGAACAGC
GCAGGGCAGCGAGAGGAGTACTTGCTGTGTTTCCACGAATTTGGAGTGTTCTGTGGATTCT
TACGGAAGACGTAGCCGCACAGACGATCTCAAGTGGAGTCGCTTACCTTTGGCCTTTGCC
TACAGAGAACCCTATCTGTTGTGACCCACTTCAACTCACTCGAAGTAATTGAGATCCAG
GCACGCTCCTCAGCAGGGACCCCTGCCCCGAGCGTACCTGGACATCCCGAACC CGCGCTAC
CTGGGCCCTGCCATTTCCCTCAGGAGCGATTTACTTGGCGTCCTCATACCAGGATAAATTA
AGGGTCATTTGCTGCAAGGGAAACCTCGTGAAGGAGTCCGGCACTGAACACCACCGGGGC
CCGTCCACCTCCCGCAGCAGCCCCAACAAGCGAGGGCCACCCACGTACAACGAGCACATC
ACCAAGCGCGTGGCCTCCAGCCCAGCGCCGCCCCGAAGGCCCCAGCCACCCGCGAGAGCCA
AGCACACCCACCGCTACCGCGAGGGGCGGACCGAGCTGCGCAGGGACAAGTCTCCTGGC
CGCCCCCTGGAGCGAGAGAAAGTCCCCCGGCCGGATGCTCAGCACGCGGAGAGAGCGGTCC
CCCGGGAGGCTGTTTGAAGACAGCAGCAGGGGCGGGCTGCCCTGCGGGAGCCGTGAGGACC
CCGCTGTCCCAGGTGAACAAGGTGAGGCAGCATTCC

- 1 -

SEQUENCE LISTING

<110> Bayer AG

<120> REGULATION OF HUMAN CITRON RHO/RAC-INTERACTING KINASE

<130> Lio 372

<150> USSN 60/301,841

<151> 2001-07-02

<150> USSN 60/338,651

<151> 2001-12-11

<150> USSN 60/375,014

<151> 2002-04-25

<160> 8

<170> PatentIn version 3.1

<210> 1

<211> 6165

<212> DNA

<213> Homo sapiens

<400> 1

atgttgaaagt tcaaatatgg agcgcggaat cctttggatg ctggtgctgc tgaaccatt	60
gccagccggg cctccaggct gaatctgttc ttccagggga aaccaccctt tatgactcaa	120
cagcagatgt ctctctttc ccgagaaggg atattagatg ccctctttgt tctctttgaa	180
gaatgcagtc agcctgctct gatgaagatt aagcacgtga gcaactttgt ccggaagtat	240
tccgacacca tagctgagtt acaggagctc cagccttcgg caaaggactt cgaagtcaga	300
agtctttag gttgtggtca ctttctgtaa gtgcagggtg taagagagaa agcaaccggg	360
gacatctatg ctatgaaagt gatgaagaag aaggctttat tggcccagga gcaggtttca	420

- 2 -

ttttttgagg aagagcggaa catattatct cgaagcacia gcccgaggat cccccaatta 480
cagtatgcct ttcaggacaa aaatcacctt tatctgggtca tggaatatca gcctggaggg 540
gaactgctgt cacttttgaa tagatatgag gaccagttag atgaaaacct gatacagttt 600
tacctagctg agctgatttt ggctgttcac agcgttcato tgatgggata cgtgcatcga 660
gacatcaagc ctgagaacat tctcgttgac cgcacaggac acatcaagct ggtggatttt 720
ggatctgccg cgaaaatgaa ttcaacaag atggtgaatg ccaaactccc gattgggacc 780
ccagattaca tggtcctga agtgctgact gtgatgaacg gggatggaaa aggcacctac 840
ggcctggact gtgactggtg gtcagtgggc gtgattgcct atgagatgat ttatgggaga 900
tcccccttcg cagagggaaac ctctgccaga accttcaata acattatgaa tttccagcgg 960
tttttgaaat ttccagatga ccccaaagtg agcagtgact ttcttgatct gattcaaagc 1020
ttgttggtcg gccagaaaga gagactgaag tttgaaggtc tttgctgcca tctttcttc 1080
totaaaattg actggaacaa cattcgtaac totcctccc ccttcgttcc caccctcaag 1140
totgacgatg acacctccaa ttttgatgaa ccagagaaga attcgtgggt ttcactctct 1200
ccgtgccagc tgagccccto aggcctctctg ggtgaagaac tgccgtttgt ggggttttcg 1260
tacagcaagg cactggggat tcttggtaga tctgagtctg ttgtgtoggg tctggactcc 1320
cctgccaaga ctagctccat ggaaaagaaa cttctcatca aaagcaaaga gctacaagac 1380
tctcaggaca agtgtcacia gatggagcag gaaatgacct gggtacatcg gagagtgtca 1440
gaggtggagg ctgtgcttag tcagaaggag gtggagctga aggcctctga gactcagaga 1500
tccctcctgg agcaggacct tgctacctac atcacagaat gcagtagcct aaagcgaagt 1560
ttggagcaag cacggatgga ggtgtcccag gaggatgaca aagcaactgca gcttctccat 1620

gatatcagag agcagagccg gaagctccaa gaaatcaaag agcaggagta ccaggctcaa 1680
gtggaagaaa tgaggttgat gatgaatcag ttggaagagg atcttgtctc agcaagaaga 1740
cggagtgatc tctacgaatc tgagctgaga gagtctcggc ttgctgctga agaattcaag 1800
cggaaagcga cagaatgtca gcataaactg ttgaaggota aggatcaagg gaagcctgaa 1860
gtgggagaat atgcgaaact ggagaagatc aatgctgagc agcagctcaa aattcaggag 1920
ctccaagaga aactggagaa ggctgtaaaa gccagcacgg aggccaccga gctgctgcag 1980
aatatccgcc aggcaaagga gcgagccgag agggagctgg agaagctgca gaaccgagag 2040
gattttcttg aaggcatcag aaagaagctg gtggaagctg aggaacgccg ccattctctg 2100
gagaacaagg taaagagact agagaccatg gagcgtagag aaaacagact gaaggatgac 2160
atccagacaa aatcccaaca gatccagcag atggctgata aaattctgga gctcgaagag 2220
aaacatcggg aggcccaagt ctgagcccag cacctagaag tgcacctgaa acagaaagag 2280
cagcactatg aggaaaagat taaagtgttg gacaatcaga taaagaaaga cctggctgac 2340
aaggagacac tggagaacat gatgcagaga cacgaggagg aggcccatga gaagggcaaa 2400
attctcagcg aacagaaggc gatgatcaat gctatggatt ccaagatcag atccctggaa 2460
cagaggattg tggaactgtc tgaagccaat aaacttgcag caaatagcag tctttttacc 2520
caaaggaaca tgaaggccca agaagagatg atttctgaac tcaggcaaca gaaattttac 2580
ctggagacac aggctgggaa gttggaggcc cagaaccgaa aactggagga gcagctggag 2640
aagatcagcc accaagacca cagtgaacag aatcggctgc tggaactgga gacaagattg 2700
cgggaggtca gtctagagca cgaggagcag aaactggagc tcaagcgcca gctcacagag 2760
ctacagctct ccctgcagga gcgcgagtca cagttgacag ccctgcaggc tgcacgggcg 2820

gccctggaga gccagcttcg ccaggcgaag acagagctgg aagagaccac agcagaagct 2880
gaagaggaga tccaggcact cacggcacat agagatgaaa tccagcgcaa atttgatgct 2940
cttcgtaaca gctgtactgt aatcacagac ctggaggagc agctaaacca gctgaccgag 3000
gacaacgctg aactcaacaa ocaaaacttc tacttgtcca aacaactcga tgaggcttct 3060
ggcgccaacg acgagattgt acaactgcga agtgaagtgg accatctccg ccgggagatc 3120
acggaacgag agatgcagct taccagccag aagcaaacga tggaggctct gaagaccacg 3180
tgcaccatgc tggaggaaca ggtcatggat ttggaggccc taaacgatga gctgctagaa 3240
aaagagcggc agtgggaggc ctggaggagc gtctgggtg atgagaaatc ccagtttgag 3300
tgtcgggttc gagagctgca gagaatgctg gacaccgaga aacagagcag ggcgagagcc 3360
gatcagcgga tcaccgagtc tcgccagggtg gtggagctgg cagtgaagga gcacaaggct 3420
gagattctcg ctctgcagca ggctctcaaa gagcagaagc tgaaggccga gagcctctct 3480
gacaagctca atgacctgga gaagaagcat gctatgcttg aaatgaatgc ccgaagctta 3540
cagcagaagc tggagactga acgagagctc aaacagaggo ttctggaaga gcaagccaaa 3600
ttacagcagc agatggacct gcagaaaaat cacattttcc gtctgactca aggactgcaa 3660
gaagctctag atcgggctga tctactgaag acagaaagaa gtgacttgga gtatcagctg 3720
gaaaacattc aggttctcta ttctcatgaa aaggtgaaaa tggaaggcac tattttctca 3780
caaaccaaac tcattgattt tctgcaagcc aaaatggacc aacctgctaa aaagaaaaag 3840
gttcctctgc agtacaatga gctgaagctg gccctggaga aggagaaagc tcgctgtgca 3900
gagctagagg aagcccttca gaagaccgc atcgagctcc ggtccgcccg ggaggaagct 3960
gcccaccgca aagcaacgga ccaccacac ccatccacgc cagccaccgc gaggcagcag 4020

atcgccatgt ccgccatcgt gcggtcgcca gagcaccagc ccagtgccat gagcctgctg 4080
gccccgccat ccagccgcag aaaggagtct tcaactccag aggaatttag tcggcgtctt 4140
aaggaacgca tgcaccacaa tattcctcac cgattcaacg taggactgaa catgcgagcc 4200
acaaagtgtg ctgtgtgtct ggataccgtg cactttggac gccaggcatc caaatgtctc 4260
gaatgtcagg tgatgtgtca cccaagtgc tccacgtgtg tgccagccac ctgcggcttg 4320
cctgtctgaat atgccacaca cttcaccgag gcctttgtgc gtgacaaaat gaactcccca 4380
ggtctccaga ccaaggagcc cagcagcagc ttgcacctgg aagggtggat gaagggtccc 4440
aggaataaca aacgaggaca gcaaggctgg gacaggaagt acattgtcct ggagggatca 4500
aaagtctca tttatgacaa tgaagccaga gaagetggac agaggccggt ggaagaattt 4560
gagctgtgcc ttccgcacgg ggatgtatct attcatggtg ccgttgggtg ttccgaactc 4620
gcaaatacag ccaaagcaga tgtcccatc atactgaaga tggaatctca ccgcacacc 4680
acctgtggc ccgggagAAC cctctacttg ctagctccca gttccctga caaacagcgc 4740
tggttcaccg ccttagaatc agttgtcgca ggtgggagag tttctaggga aaaagcagaa 4800
gctgatgcta aactgcttgg aaactccctg ctgaaactgg aaggtgatga ccgtctagac 4860
atgaactgca cgtgcctt cagtgaccag gtggtgttgg tgggcaccga ggaagggtc 4920
tacgcctga atgtcttgaa aaactcccta acccatgtcc caggaaattgg agcagtcttc 4980
caaatttata ttatcaagga cctggagaag ctactcatga tagcaggaga agagcgggca 5040
ctgtgtcttg tggacgtgaa gaaagtgaaa cagtccctgg ccagtcacca cctgcctgcc 5100
cagcccgaca tctcaccacaa catttttgaa gctgtcaagg gctgccactt gtttggggca 5160
ggcaagattg agaacgggct ctgcatctgt gcagccatgc ccagcaaagt cgtcattctc 5220

- 6 -

cgctacaacg aaaacctcag caaatactgc atccggaaag agatagagac ctcagagccc 5280
tgcagctgta tccacttcac caattacagt atcctcattg gaaccaataa attctacgaa 5340
atcgacatga agcagtacac gctcgaggaa ttcttgata agaatgacca ttccttgga 5400
cctgctgtgt ttgcccctc ttccaacagc ttccctgtct caatcgtgca ggtgaacagc 5460
gcagggcagc gagaggagta cttgctgtgt ttccacgaat ttggagtgtt cgtggattct 5520
tacggaagac gtagccgcac agacgatctc aagtggagtc gcttaccttt ggcctttgcc 5580
tacagagaac cctatctgtt tgtgaccac ttcaactcac tcgaagtaat tgagatccag 5640
gcacgctcct cagcaggagc ccctgcccga gcgtagctg acatcccga cccgcgctac 5700
ctgggcccctg ccatttctc aggagcgatt taottggcgt cctcatacca ggataaatta 5760
agggtcattt gctgcaaggg aaacctcgtg aaggagtccg gcaactgaaca ccaccggggc 5820
ccgtccacct ccgcagcag cccaacaag cgaggccac ccacgtacaa cgagcacatc 5880
accaagcgcg tggcctccag ccagcgccg ccgaaggcc ccagccacc gcgagagcca 5940
agcacacccc accgctaccg cgaggggagg accgagctgc gcagggacaa gtctcctggc 6000
cgccccctgg agcgagagaa gtccccggc cggatgctca gcacgggag agagcggtcc 6060
cccgggaggc tgtttgaaga cagcagcagg ggccggctgc ctgcccggag cgtgaggacc 6120
ccgtgtccc aggtgaacaa ggtctgggac cagtcttcag tataa 6165

<210> 2

<211> 2054

<212> PRT

<213> Homo sapiens

<400> 2

- 7 -

Met Leu Lys Phe Lys Tyr Gly Ala Arg Asn Pro Leu Asp Ala Gly Ala
 1 5 10 15

Ala Glu Pro Ile Ala Ser Arg Ala Ser Arg Leu Asn Leu Phe Phe Gln
 20 25 30

Gly Lys Pro Pro Phe Met Thr Gln Gln Gln Met Ser Pro Leu Ser Arg
 35 40 45

Glu Gly Ile Leu Asp Ala Leu Phe Val Leu Phe Glu Glu Cys Ser Gln
 50 55 60

Pro Ala Leu Met Lys Ile Lys His Val Ser Asn Phe Val Arg Lys Tyr
 65 70 75 80

Ser Asp Thr Ile Ala Glu Leu Gln Glu Leu Gln Pro Ser Ala Lys Asp
 85 90 95

Phe Glu Val Arg Ser Leu Val Gly Cys Gly His Phe Ala Glu Val Gln
 100 105 110

Val Val Arg Glu Lys Ala Thr Gly Asp Ile Tyr Ala Met Lys Val Met
 115 120 125

Lys Lys Lys Ala Leu Leu Ala Gln Glu Gln Val Ser Phe Phe Glu Glu
 130 135 140

Glu Arg Asn Ile Leu Ser Arg Ser Thr Ser Pro Trp Ile Pro Gln Leu
 145 150 155 160

- 8 -

Gln Tyr Ala Phe Gln Asp Lys Asn His Leu Tyr Leu Val Met Glu Tyr
165 170 175

Gln Pro Gly Gly Asp Leu Leu Ser Leu Leu Asn Arg Tyr Glu Asp Gln
180 185 190

Leu Asp Glu Asn Leu Ile Gln Phe Tyr Leu Ala Glu Leu Ile Leu Ala
195 200 205

Val His Ser Val His Leu Met Gly Tyr Val His Arg Asp Ile Lys Pro
210 215 220

Glu Asn Ile Leu Val Asp Arg Thr Gly His Ile Lys Leu Val Asp Phe
225 230 235 240

Gly Ser Ala Ala Lys Met Asn Ser Asn Lys Met Val Asn Ala Lys Leu
245 250 255

Pro Ile Gly Thr Pro Asp Tyr Met Ala Pro Glu Val Leu Thr Val Met
260 265 270

Asn Gly Asp Gly Lys Gly Thr Tyr Gly Leu Asp Cys Asp Trp Trp Ser
275 280 285

Val Gly Val Ile Ala Tyr Glu Met Ile Tyr Gly Arg Ser Pro Phe Ala
290 295 300

Glu Gly Thr Ser Ala Arg Thr Phe Asn Asn Ile Met Asn Phe Gln Arg
305 310 315 320

- 9 -

Phe Leu Lys Phe Pro Asp Asp Pro Lys Val Ser Ser Asp Phe Leu Asp
325 330 335

Leu Ile Gln Ser Leu Leu Cys Gly Gln Lys Glu Arg Leu Lys Phe Glu
340 345 350

Gly Leu Cys Cys His Pro Phe Phe Ser Lys Ile Asp Trp Asn Asn Ile
355 360 365

Arg Asn Ser Pro Pro Pro Phe Val Pro Thr Leu Lys Ser Asp Asp Asp
370 375 380

Thr Ser Asn Phe Asp Glu Pro Glu Lys Asn Ser Trp Val Ser Ser Ser
385 390 395 400

Pro Cys Gln Leu Ser Pro Ser Gly Phe Ser Gly Glu Glu Leu Pro Phe
405 410 415

Val Gly Phe Ser Tyr Ser Lys Ala Leu Gly Ile Leu Gly Arg Ser Glu
420 425 430

Ser Val Val Ser Gly Leu Asp Ser Pro Ala Lys Thr Ser Ser Met Glu
435 440 445

Lys Lys Leu Leu Ile Lys Ser Lys Glu Leu Gln Asp Ser Gln Asp Lys
450 455 460

Cys His Lys Met Glu Gln Glu Met Thr Arg Leu His Arg Arg Val Ser
465 470 475 480

- 10 -

Glu Val Glu Ala Val Leu Ser Gln Lys Glu Val Glu Leu Lys Ala Ser
485 490 495

Glu Thr Gln Arg Ser Leu Leu Glu Gln Asp Leu Ala Thr Tyr Ile Thr
500 505 510

Glu Cys Ser Ser Leu Lys Arg Ser Leu Glu Gln Ala Arg Met Glu Val
515 520 525

Ser Gln Glu Asp Asp Lys Ala Leu Gln Leu Leu His Asp Ile Arg Glu
530 535 540

Gln Ser Arg Lys Leu Gln Glu Ile Lys Glu Gln Glu Tyr Gln Ala Gln
545 550 555 560

Val Glu Glu Met Arg Leu Met Met Asn Gln Leu Glu Glu Asp Leu Val
565 570 575

Ser Ala Arg Arg Arg Ser Asp Leu Tyr Glu Ser Glu Leu Arg Glu Ser
580 585 590

Arg Leu Ala Ala Glu Glu Phe Lys Arg Lys Ala Thr Glu Cys Gln His
595 600 605

Lys Leu Leu Lys Ala Lys Asp Gln Gly Lys Pro Glu Val Gly Glu Tyr
610 615 620

Ala Lys Leu Glu Lys Ile Asn Ala Glu Gln Gln Leu Lys Ile Gln Glu
625 630 635 640

- 11 -

Leu Gln Glu Lys Leu Glu Lys Ala Val Lys Ala Ser Thr Glu Ala Thr
645 650 655

Glu Leu Leu Gln Asn Ile Arg Gln Ala Lys Glu Arg Ala Glu Arg Glu
660 665 670

Leu Glu Lys Leu Gln Asn Arg Glu Asp Ser Ser Glu Gly Ile Arg Lys
675 680 685

Lys Leu Val Glu Ala Glu Glu Arg Arg His Ser Leu Glu Asn Lys Val
690 695 700

Lys Arg Leu Glu Thr Met Glu Arg Arg Glu Asn Arg Leu Lys Asp Asp
705 710 715 720

Ile Gln Thr Lys Ser Gln Gln Ile Gln Gln Met Ala Asp Lys Ile Leu
725 730 735

Glu Leu Glu Glu Lys His Arg Glu Ala Gln Val Ser Ala Gln His Leu
740 745 750

Glu Val His Leu Lys Gln Lys Glu Gln His Tyr Glu Glu Lys Ile Lys
755 760 765

Val Leu Asp Asn Gln Ile Lys Lys Asp Leu Ala Asp Lys Glu Thr Leu
770 775 780

Glu Asn Met Met Gln Arg His Glu Glu Glu Ala His Glu Lys Gly Lys
785 790 795 800

- 12 -

Ile Leu Ser Glu Gln Lys Ala Met Ile Asn Ala Met Asp Ser Lys Ile
805 810 815

Arg Ser Leu Glu Gln Arg Ile Val Glu Leu Ser Glu Ala Asn Lys Leu
820 825 830

Ala Ala Asn Ser Ser Leu Phe Thr Gln Arg Asn Met Lys Ala Gln Glu
835 840 845

Glu Met Ile Ser Glu Leu Arg Gln Gln Lys Phe Tyr Leu Glu Thr Gln
850 855 860

Ala Gly Lys Leu Glu Ala Gln Asn Arg Lys Leu Glu Glu Gln Leu Glu
865 870 875 880

Lys Ile Ser His Gln Asp His Ser Asp Lys Asn Arg Leu Leu Glu Leu
885 890 895

Glu Thr Arg Leu Arg Glu Val Ser Leu Glu His Glu Glu Gln Lys Leu
900 905 910

Glu Leu Lys Arg Gln Leu Thr Glu Leu Gln Leu Ser Leu Gln Glu Arg
915 920 925

Glu Ser Gln Leu Thr Ala Leu Gln Ala Ala Arg Ala Ala Leu Glu Ser
930 935 940

Gln Leu Arg Gln Ala Lys Thr Glu Leu Glu Glu Thr Thr Ala Glu Ala
945 950 955 960

- 13 -

Glu Glu Glu Ile Gln Ala Leu Thr Ala His Arg Asp Glu Ile Gln Arg
965 970 975

Lys Phe Asp Ala Leu Arg Asn Ser Cys Thr Val Ile Thr Asp Leu Glu
980 985 990

Glu Gln Leu Asn Gln Leu Thr Glu Asp Asn Ala Glu Leu Asn Asn Gln
995 1000 1005

Asn Phe Tyr Leu Ser Lys Gln Leu Asp Glu Ala Ser Gly Ala Asn
1010 1015 1020

Asp Glu Ile Val Gln Leu Arg Ser Glu Val Asp His Leu Arg Arg
1025 1030 1035

Glu Ile Thr Glu Arg Glu Met Gln Leu Thr Ser Gln Lys Gln Thr
1040 1045 1050

Met Glu Ala Leu Lys Thr Thr Cys Thr Met Leu Glu Glu Gln Val
1055 1060 1065

Met Asp Leu Glu Ala Leu Asn Asp Glu Leu Leu Glu Lys Glu Arg
1070 1075 1080

Gln Trp Glu Ala Trp Arg Ser Val Leu Gly Asp Glu Lys Ser Gln
1085 1090 1095

Phe Glu Cys Arg Val Arg Glu Leu Gln Arg Met Leu Asp Thr Glu
1100 1105 1110

- 14 -

Lys Gln Ser Arg Ala Arg Ala Asp Gln Arg Ile Thr Glu Ser Arg
1115 1120 1125

Gln Val Val Glu Leu Ala Val Lys Glu His Lys Ala Glu Ile Leu
1130 1135 1140

Ala Leu Gln Gln Ala Leu Lys Glu Gln Lys Leu Lys Ala Glu Ser
1145 1150 1155

Leu Ser Asp Lys Leu Asn Asp Leu Glu Lys Lys His Ala Met Leu
1160 1165 1170

Glu Met Asn Ala Arg Ser Leu Gln Gln Lys Leu Glu Thr Glu Arg
1175 1180 1185

Glu Leu Lys Gln Arg Leu Leu Glu Glu Gln Ala Lys Leu Gln Gln
1190 1195 1200

Gln Met Asp Leu Gln Lys Asn His Ile Phe Arg Leu Thr Gln Gly
1205 1210 1215

Leu Gln Glu Ala Leu Asp Arg Ala Asp Leu Leu Lys Thr Glu Arg
1220 1225 1230

Ser Asp Leu Glu Tyr Gln Leu Glu Asn Ile Gln Val Leu Tyr Ser
1235 1240 1245

His Glu Lys Val Lys Met Glu Gly Thr Ile Ser Gln Gln Thr Lys
1250 1255 1260

- 15 -

Leu Ile Asp Phe Leu Gln Ala Lys Met Asp Gln Pro Ala Lys Lys
 1265 1270 1275

Lys Lys Val Pro Leu Gln Tyr Asn Glu Leu Lys Leu Ala Leu Glu
 1280 1285 1290

Lys Glu Lys Ala Arg Cys Ala Glu Leu Glu Glu Ala Leu Gln Lys
 1295 1300 1305

Thr Arg Ile Glu Leu Arg Ser Ala Arg Glu Glu Ala Ala His Arg
 1310 1315 1320

Lys Ala Thr Asp His Pro His Pro Ser Thr Pro Ala Thr Ala Arg
 1325 1330 1335

Gln Gln Ile Ala Met Ser Ala Ile Val Arg Ser Pro Glu His Gln
 1340 1345 1350

Pro Ser Ala Met Ser Leu Leu Ala Pro Pro Ser Ser Arg Arg Lys
 1355 1360 1365

Glu Ser Ser Thr Pro Glu Glu Phe Ser Arg Arg Leu Lys Glu Arg
 1370 1375 1380

Met His His Asn Ile Pro His Arg Phe Asn Val Gly Leu Asn Met
 1385 1390 1395

Arg Ala Thr Lys Cys Ala Val Cys Leu Asp Thr Val His Phe Gly
 1400 1405 1410

- 16 -

Arg Gln Ala Ser Lys Cys Leu Glu Cys Gln Val Met Cys His Pro
1415 1420 1425

Lys Cys Ser Thr Cys Leu Pro Ala Thr Cys Gly Leu Pro Ala Glu
1430 1435 1440

Tyr Ala Thr His Phe Thr Glu Ala Phe Cys Arg Asp Lys Met Asn
1445 1450 1455

Ser Pro Gly Leu Gln Thr Lys Glu Pro Ser Ser Ser Leu His Leu
1460 1465 1470

Glu Gly Trp Met Lys Val Pro Arg Asn Asn Lys Arg Gly Gln Gln
1475 1480 1485

Gly Trp Asp Arg Lys Tyr Ile Val Leu Glu Gly Ser Lys Val Leu
1490 1495 1500

Ile Tyr Asp Asn Glu Ala Arg Glu Ala Gly Gln Arg Pro Val Glu
1505 1510 1515

Glu Phe Glu Leu Cys Leu Pro Asp Gly Asp Val Ser Ile His Gly
1520 1525 1530

Ala Val Gly Ala Ser Glu Leu Ala Asn Thr Ala Lys Ala Asp Val
1535 1540 1545

Pro Tyr Ile Leu Lys Met Glu Ser His Pro His Thr Thr Cys Trp
1550 1555 1560

- 17 -

Pro Gly	Arg Thr	Leu Tyr	Leu	Leu Ala	Pro Ser	Phe	Pro Asp	Lys
1565			1570			1575		

Gln Arg	Trp Val	Thr Ala	Leu	Glu Ser	Val Val	Ala	Gly Gly	Arg
1580			1585			1590		

Val Ser	Arg Glu	Lys Ala	Glu	Ala Asp	Ala Lys	Leu	Leu Gly	Asn
1595			1600			1605		

Ser Leu	Leu Lys	Leu Glu	Gly	Asp Asp	Arg Leu	Asp	Met Asn	Cys
1610			1615			1620		

Thr Leu	Pro Phe	Ser Asp	Gln	Val Val	Leu Val	Gly	Thr Glu	Glu
1625			1630			1635		

Gly Leu	Tyr Ala	Leu Asn	Val	Leu Lys	Asn Ser	Leu	Thr His	Val
1640			1645			1650		

Pro Gly	Ile Gly	Ala Val	Phe	Gln Ile	Tyr Ile	Ile	Lys Asp	Leu
1655			1660			1665		

Glu Lys	Leu Leu	Met Ile	Ala	Gly Glu	Glu Arg	Ala	Leu Cys	Leu
1670			1675			1680		

Val Asp	Val Lys	Lys Val	Lys	Gln Ser	Leu Ala	Gln	Ser His	Leu
1685			1690			1695		

Pro Ala	Gln Pro	Asp Ile	Ser	Pro Asn	Ile Phe	Glu	Ala Val	Lys
1700			1705			1710		

- 18 -

Gly Cys His Leu Phe Gly Ala Gly Lys Ile Glu Asn Gly Leu Cys
1715 1720 1725

Ile Cys Ala Ala Met Pro Ser Lys Val Val Ile Leu Arg Tyr Asn
1730 1735 1740

Glu Asn Leu Ser Lys Tyr Cys Ile Arg Lys Glu Ile Glu Thr Ser
1745 1750 1755

Glu Pro Cys Ser Cys Ile His Phe Thr Asn Tyr Ser Ile Leu Ile
1760 1765 1770

Gly Thr Asn Lys Phe Tyr Glu Ile Asp Met Lys Gln Tyr Thr Leu
1775 1780 1785

Glu Glu Phe Leu Asp Lys Asn Asp His Ser Leu Ala Pro Ala Val
1790 1795 1800

Phe Ala Ala Ser Ser Asn Ser Phe Pro Val Ser Ile Val Gln Val
1805 1810 1815

Asn Ser Ala Gly Gln Arg Glu Glu Tyr Leu Leu Cys Phe His Glu
1820 1825 1830

Phe Gly Val Phe Val Asp Ser Tyr Gly Arg Arg Ser Arg Thr Asp
1835 1840 1845

Asp Leu Lys Trp Ser Arg Leu Pro Leu Ala Phe Ala Tyr Arg Glu
1850 1855 1860

- 19 -

Pro Tyr Leu Phe Val Thr His Phe Asn Ser Leu Glu Val Ile Glu
 1865 1870 1875

Ile Gln Ala Arg Ser Ser Ala Gly Thr Pro Ala Arg Ala Tyr Leu
 1880 1885 1890

Asp Ile Pro Asn Pro Arg Tyr Leu Gly Pro Ala Ile Ser Ser Gly
 1895 1900 1905

Ala Ile Tyr Leu Ala Ser Ser Tyr Gln Asp Lys Leu Arg Val Ile
 1910 1915 1920

Cys Cys Lys Gly Asn Leu Val Lys Glu Ser Gly Thr Glu His His
 1925 1930 1935

Arg Gly Pro Ser Thr Ser Arg Ser Ser Pro Asn Lys Arg Gly Pro
 1940 1945 1950

Pro Thr Tyr Asn Glu His Ile Thr Lys Arg Val Ala Ser Ser Pro
 1955 1960 1965

Ala Pro Pro Glu Gly Pro Ser His Pro Arg Glu Pro Ser Thr Pro
 1970 1975 1980

His Arg Tyr Arg Glu Gly Arg Thr Glu Leu Arg Arg Asp Lys Ser
 1985 1990 1995

Pro Gly Arg Pro Leu Glu Arg Glu Lys Ser Pro Gly Arg Met Leu
 2000 2005 2010

- 20 -

Ser Thr Arg Arg Glu Arg Ser Pro Gly Arg Leu Phe Glu Asp Ser
 2015 2020 2025

Ser Arg Gly Arg Leu Pro Ala Gly Ala Val Arg Thr Pro Leu Ser
 2030 2035 2040

Gln Val Asn Lys Val Trp Asp Gln Ser Ser Val
 2045 2050

<210> 3

<211> 2055

<212> PRT

<213> Homo sapiens

<400> 3

Met Leu Lys Phe Lys Tyr Gly Val Arg Asn Pro Pro Glu Ala Ser Ala
 1 5 10 15

Ser Glu Pro Ile Ala Ser Arg Ala Ser Arg Leu Asn Leu Phe Phe Gln
 20 25 30

Gly Lys Pro Pro Leu Met Thr Gln Gln Met Ser Ala Leu Ser Arg
 35 40 45

Glu Gly Met Leu Asp Ala Leu Phe Ala Leu Phe Glu Glu Cys Ser Gln
 50 55 60

Pro Ala Leu Met Lys Met Lys His Val Ser Ser Phe Val Gln Lys Tyr
 65 70 75 80

- 21 -

Ser Asp Thr Ile Ala Glu Leu Arg Glu Leu Gln Pro Ser Ala Arg Asp
85 90 95

Phe Glu Val Arg Ser Leu Val Gly Cys Gly His Phe Ala Glu Val Gln
100 105 110

Val Val Arg Glu Lys Ala Thr Gly Asp Val Tyr Ala Met Lys Ile Met
115 120 125

Lys Lys Lys Ala Leu Leu Ala Gln Glu Gln Val Ser Phe Phe Glu Glu
130 135 140

Glu Arg Asn Ile Leu Ser Arg Ser Thr Ser Pro Trp Ile Pro Gln Leu
145 150 155 160

Gln Tyr Ala Phe Gln Asp Lys Asn Asn Leu Tyr Leu Val Met Glu Tyr
165 170 175

Gln Pro Gly Gly Asp Phe Leu Ser Leu Leu Asn Arg Tyr Glu Asp Gln
180 185 190

Leu Asp Glu Ser Met Ile Gln Phe Tyr Leu Ala Glu Leu Ile Leu Ala
195 200 205

Val His Ser Val His Gln Met Gly Tyr Val His Arg Asp Ile Lys Pro
210 215 220

Glu Asn Ile Leu Ile Asp Arg Thr Gly Glu Ile Lys Leu Val Asp Phe
225 230 235 240

- 22 -

Gly Ser Ala Ala Lys Met Asn Ser Asn Lys Val Asp Ala Lys Leu Pro
245 250 255

Ile Gly Thr Pro Asp Tyr Met Ala Pro Glu Val Leu Thr Val Met Asn
260 265 270

Glu Asp Arg Arg Gly Thr Tyr Gly Leu Asp Cys Asp Trp Trp Ser Val
275 280 285

Gly Val Val Ala Tyr Glu Met Val Tyr Gly Lys Thr Pro Phe Thr Glu
290 295 300

Gly Thr Ser Ala Arg Thr Phe Asn Asn Ile Met Asn Phe Gln Arg Phe
305 310 315 320

Leu Lys Phe Pro Asp Asp Pro Lys Val Ser Ser Glu Leu Leu Asp Leu
325 330 335

Leu Gln Ser Leu Leu Cys Val Gln Lys Glu Arg Leu Lys Phe Glu Gly
340 345 350

Leu Cys Cys His Pro Phe Phe Ala Arg Thr Asp Trp Asn Asn Ile Arg
355 360 365

Asn Ser Pro Pro Pro Phe Val Pro Thr Leu Lys Ser Asp Asp Asp Thr
370 375 380

Ser Asn Phe Asp Glu Pro Glu Lys Asn Ser Trp Ala Phe Ile Leu Cys
385 390 395 400

- 23 -

Val Pro Ala Glu Pro Leu Ala Phe Ser Gly Glu Glu Leu Pro Phe Val
405 410 415

Gly Phe Ser Tyr Ser Lys Ala Leu Gly Tyr Leu Gly Arg Ser Glu Ser
420 425 430

Val Val Ser Ser Leu Asp Ser Pro Ala Lys Val Ser Ser Met Glu Lys
435 440 445

Lys Leu Leu Ile Lys Ser Lys Glu Leu Gln Asp Ser Gln Asp Lys Cys
450 455 460

His Lys Met Glu Gln Glu Met Thr Arg Leu His Arg Arg Val Ser Glu
465 470 475 480

Val Glu Ala Val Leu Ser Gln Lys Glu Val Glu Leu Lys Ala Ser Glu
485 490 495

Thr Gln Arg Ser Leu Leu Glu Gln Asp Leu Ala Thr Tyr Ile Thr Glu
500 505 510

Cys Ser Ser Leu Lys Arg Ser Leu Glu Gln Ala Arg Met Glu Val Ser
515 520 525

Gln Glu Asp Asp Lys Ala Leu Gln Leu Leu His Asp Ile Arg Glu Gln
530 535 540

Ser Arg Lys Leu Gln Glu Ile Lys Glu Gln Glu Tyr Gln Ala Gln Val
545 550 555 560

- 24 -

Glu Glu Met Arg Leu Met Met Asn Gln Leu Glu Glu Asp Leu Val Ser
565 570 575

Ala Arg Arg Arg Ser Asp Leu Tyr Glu Ser Glu Leu Arg Glu Ser Arg
580 585 590

Leu Ala Ala Glu Glu Phe Lys Arg Lys Ala Asn Glu Cys Gln His Lys
595 600 605

Leu Met Lys Ala Lys Asp Gln Gly Lys Pro Glu Val Gly Glu Tyr Ser
610 615 620

Lys Leu Glu Lys Ile Asn Ala Glu Gln Gln Leu Lys Ile Gln Glu Leu
625 630 635 640

Gln Glu Lys Leu Glu Lys Ala Val Lys Ala Ser Thr Glu Ala Thr Glu
645 650 655

Leu Leu Gln Asn Ile Arg Gln Ala Lys Glu Arg Ala Glu Arg Glu Leu
660 665 670

Glu Lys Leu His Asn Arg Glu Asp Ser Ser Glu Gly Ile Lys Lys Lys
675 680 685

Leu Val Glu Ala Glu Glu Arg Arg His Ser Leu Glu Asn Lys Val Lys
690 695 700

Arg Leu Glu Thr Met Glu Arg Arg Glu Asn Arg Leu Lys Asp Asp Ile
705 710 715 720

- 25 -

Gln Thr Lys Ser Glu Gln Ile Gln Gln Met Ala Asp Lys Ile Leu Glu
725 730 735

Leu Glu Glu Lys His Arg Glu Ala Gln Val Ser Ala Gln His Leu Glu
740 745 750

Val His Leu Lys Gln Lys Glu Gln His Tyr Glu Glu Lys Ile Lys Val
755 760 765

Leu Asp Asn Gln Ile Lys Lys Asp Leu Ala Asp Lys Glu Ser Leu Glu
770 775 780

Asn Met Met Gln Arg His Glu Glu Glu Ala His Glu Lys Gly Lys Ile
785 790 795 800

Leu Ser Glu Gln Lys Ala Met Ile Asn Ala Met Asp Ser Lys Ile Arg
805 810 815

Ser Leu Glu Gln Arg Ile Val Glu Leu Ser Glu Ala Asn Lys Leu Ala
820 825 830

Ala Asn Ser Ser Leu Phe Thr Gln Arg Asn Met Lys Ala Gln Glu Glu
835 840 845

Met Ile Ser Glu Leu Arg Gln Gln Lys Phe Tyr Leu Glu Thr Gln Ala
850 855 860

Gly Lys Leu Glu Ala Gln Asn Arg Lys Leu Glu Glu Gln Leu Glu Lys
865 870 875 880

- 26 -

Ile Ser His Gln Asp His Ser Asp Lys Ser Arg Leu Leu Glu Leu Glu
885 890 895

Thr Arg Leu Arg Glu Val Ser Leu Glu His Glu Glu Gln Lys Leu Glu
900 905 910

Leu Lys Arg Gln Leu Thr Glu Leu Gln Leu Ser Leu Gln Glu Arg Glu
915 920 925

Ser Gln Leu Thr Ala Leu Gln Ala Ala Arg Ala Ala Leu Glu Ser Gln
930 935 940

Leu Arg Gln Ala Lys Thr Glu Leu Glu Glu Thr Thr Ala Glu Ala Glu
945 950 955 960

Glu Glu Ile Gln Ala Leu Thr Ala His Arg Asp Glu Ile Gln Arg Lys
965 970 975

Phe Asp Ala Leu Arg Asn Ser Cys Thr Val Ile Thr Asp Leu Glu Glu
980 985 990

Gln Leu Asn Gln Leu Thr Glu Asp Asn Ala Glu Leu Asn Asn Gln Asn
995 1000 1005

Phe Tyr Leu Ser Lys Gln Leu Asp Glu Ala Ser Gly Ala Asn Asp
1010 1015 1020

Glu Ile Val Gln Leu Arg Ser Glu Val Asp His Leu Arg Arg Glu
1025 1030 1035

- 27 -

Ile Thr	Glu Arg	Glu Met	Gln	Leu Thr	Ser	Gln Lys	Gln Thr Met
1040			1045			1050	

Glu Ala	Leu Lys	Thr Thr	Cys	Thr Met	Leu Glu	Glu	Gln Val Leu
1055			1060			1065	

Asp Leu	Glu Ala	Leu Asn	Asp	Glu Leu	Leu Glu	Lys	Glu Arg Gln
1070			1075			1080	

Trp Glu	Ala Trp	Arg Ser	Val	Leu Gly	Asp Glu	Lys	Ser Gln Phe
1085			1090			1095	

Glu Cys	Arg Val	Arg Glu	Leu	Gln Arg	Met Leu	Asp	Thr Glu Lys
1100			1105			1110	

Gln Ser	Arg Ala	Arg Ala	Asp	Gln Arg	Ile Thr	Glu	Ser Arg Gln
1115			1120			1125	

Val Val	Glu Leu	Ala Val	Lys	Glu His	Lys Ala	Glu	Ile Leu Ala
1130			1135			1140	

Leu Gln	Gln Ala	Leu Lys	Glu	Gln Lys	Leu Lys	Ala	Glu Ser Leu
1145			1150			1155	

Ser Asp	Lys Leu	Asn Asp	Leu	Glu Lys	Lys His	Ala	Met Leu Glu
1160			1165			1170	

Met Asn	Ala Arg	Ser Leu	Gln	Gln Lys	Leu Glu	Thr	Glu Arg Glu
1175			1180			1185	

- 28 -

Leu Lys Gln Arg Leu Leu Glu Glu Gln Ala Lys Leu Gln Gln Gln
1190 1195 1200

Met Asp Leu Gln Lys Asn His Ile Phe Arg Leu Thr Gln Gly Leu
1205 1210 1215

Gln Glu Ala Leu Asp Arg Ala Asp Leu Leu Lys Thr Glu Arg Ser
1220 1225 1230

Asp Leu Glu Tyr Gln Leu Glu Asn Ile Gln Val Leu Tyr Ser His
1235 1240 1245

Glu Lys Val Lys Met Glu Gly Thr Ile Ser Gln Gln Thr Lys Leu
1250 1255 1260

Ile Asp Phe Leu Gln Ala Lys Met Asp Gln Pro Ala Lys Lys Lys
1265 1270 1275

Lys Val Pro Leu Gln Tyr Asn Glu Leu Lys Leu Ala Leu Glu Lys
1280 1285 1290

Glu Lys Ala Arg Cys Ala Glu Leu Glu Glu Ala Leu Gln Lys Thr
1295 1300 1305

Arg Ile Glu Leu Arg Ser Ala Arg Glu Glu Ala Ala His Arg Lys
1310 1315 1320

Ala Thr Asp His Pro His Pro Ser Thr Pro Ala Thr Ala Arg Gln
1325 1330 1335

- 29 -

Gln Ile Ala Met Ser Ala Ile Val Arg Ser Pro Glu His Gln Pro
1340 1345 1350

Ser Ala Met Ser Leu Leu Ala Pro Pro Ser Ser Arg Arg Lys Glu
1355 1360 1365

Ser Ser Thr Pro Glu Glu Phe Ser Arg Arg Leu Lys Glu Arg Met
1370 1375 1380

His His Asn Ile Pro His Arg Phe Asn Val Gly Leu Asn Met Arg
1385 1390 1395

Ala Thr Lys Cys Ala Val Cys Leu Asp Thr Val His Phe Gly Arg
1400 1405 1410

Gln Ala Ser Lys Cys Leu Glu Cys Gln Val Met Cys His Pro Lys
1415 1420 1425

Cys Ser Thr Cys Leu Pro Ala Thr Cys Gly Leu Pro Ala Glu Tyr
1430 1435 1440

Ala Thr His Phe Thr Glu Ala Phe Cys Arg Asp Lys Met Asn Ser
1445 1450 1455

Pro Gly Leu Gln Ser Lys Glu Pro Gly Ser Ser Leu His Leu Glu
1460 1465 1470

Gly Trp Met Lys Val Pro Arg Asn Asn Lys Arg Gly Gln Gln Gly
1475 1480 1485

- 30 -

Trp Asp Arg Lys Tyr Ile Val Leu Glu Gly Ser Lys Val Leu Ile
1490 1495 1500

Tyr Asp Asn Glu Ala Arg Glu Ala Gly Gln Arg Pro Val Glu Glu
1505 1510 1515

Phe Glu Leu Cys Leu Pro Asp Gly Asp Val Ser Ile His Gly Ala
1520 1525 1530

Val Gly Ala Ser Glu Leu Ala Asn Thr Ala Lys Ala Asp Val Pro
1535 1540 1545

Tyr Ile Leu Lys Met Glu Ser His Pro His Thr Thr Cys Trp Pro
1550 1555 1560

Gly Arg Thr Leu Tyr Leu Leu Ala Pro Ser Phe Pro Asp Lys Gln
1565 1570 1575

Arg Trp Val Thr Ala Leu Glu Ser Val Val Ala Gly Gly Arg Val
1580 1585 1590

Ser Arg Glu Lys Ala Glu Ala Asp Ala Lys Leu Leu Gly Asn Ser
1595 1600 1605

Leu Leu Lys Leu Glu Gly Asp Asp Arg Leu Asp Met Asn Cys Thr
1610 1615 1620

Leu Pro Phe Ser Asp Gln Val Val Leu Val Gly Thr Glu Glu Gly
1625 1630 1635

- 31 -

Leu Tyr Ala Leu Asn Val Leu Lys Asn Ser Leu Thr His Ile Pro
1640 1645 1650

Gly Ile Gly Ala Val Phe Gln Ile Tyr Ile Ile Lys Asp Leu Glu
1655 1660 1665

Lys Leu Leu Met Ile Ala Gly Glu Glu Arg Ala Leu Cys Leu Val
1670 1675 1680

Asp Val Lys Lys Val Lys Gln Ser Leu Ala Gln Ser His Leu Pro
1685 1690 1695

Ala Gln Pro Asp Val Ser Pro Asn Ile Phe Glu Ala Val Lys Gly
1700 1705 1710

Cys His Leu Phe Ala Ala Gly Lys Ile Glu Asn Ser Leu Cys Ile
1715 1720 1725

Cys Ala Ala Met Pro Ser Lys Val Val Ile Leu Arg Tyr Asn Asp
1730 1735 1740

Asn Leu Ser Lys Tyr Cys Ile Arg Lys Glu Ile Glu Thr Ser Glu
1745 1750 1755

Pro Cys Ser Cys Ile His Phe Thr Asn Tyr Ser Ile Leu Ile Gly
1760 1765 1770

Thr Asn Lys Phe Tyr Glu Ile Asp Met Lys Gln Tyr Thr Leu Asp
1775 1780 1785

- 32 -

Glu Phe	Leu Asp Lys Asn Asp	His Ser Leu Ala Pro	Ala Val Phe
1790	1795	1800	

Ala Ser	Ser Ser Asn Ser Phe	Pro Val Ser Ile Val	Gln Ala Asn
1805	1810	1815	

Ser Ala	Gly Gln Arg Glu Glu	Tyr Leu Leu Cys Phe	His Glu Phe
1820	1825	1830	

Gly Val	Phe Val Asp Ser Tyr	Gly Arg Arg Ser Arg	Thr Asp Asp
1835	1840	1845	

Leu Lys	Trp Ser Arg Leu Pro	Leu Ala Phe Ala Tyr	Arg Glu Pro
1850	1855	1860	

Tyr Leu	Phe Val Thr His Phe	Asn Ser Leu Glu Val	Ile Glu Ile
1865	1870	1875	

Gln Ala	Arg Ser Ser Leu Gly	Ser Pro Ala Arg Ala	Tyr Leu Glu
1880	1885	1890	

Ile Pro	Asn Pro Arg Tyr Leu	Gly Pro Ala Ile Ser	Ser Gly Ala
1895	1900	1905	

Ile Tyr	Leu Ala Ser Ser Tyr	Gln Asp Lys Leu Arg	Val Ile Cys
1910	1915	1920	

Cys Lys	Gly Asn Leu Val Lys	Glu Ser Gly Thr Glu	Gln His Arg
1925	1930	1935	

- 33 -

Val Pro Ser Thr Ser Arg Ser Ser Pro Asn Lys Arg Gly Pro Pro
 1940 1945 1950

Thr Tyr Asn Glu His Ile Thr Lys Arg Val Ala Ser Ser Pro Ala
 1955 1960 1965

Pro Pro Glu Gly Pro Ser His Pro Arg Glu Pro Ser Thr Pro His
 1970 1975 1980

Arg Tyr Arg Asp Arg Glu Gly Arg Thr Glu Leu Arg Arg Asp Lys
 1985 1990 1995

Ser Pro Gly Arg Pro Leu Glu Arg Glu Lys Ser Pro Gly Arg Met
 2000 2005 2010

Leu Ser Thr Arg Arg Glu Arg Ser Pro Gly Arg Leu Phe Glu Asp
 2015 2020 2025

Ser Ser Arg Gly Arg Leu Pro Ala Gly Ala Val Arg Thr Pro Leu
 2030 2035 2040

Ser Gln Val Asn Lys Val Trp Asp Gln Ser Ser Val
 2045 2050 2055

<210> 4

<211> 8603

<212> DNA

<213> Homo sapiens

<400> 4

atgttgaagt tcaaatatgg agcgcggaat cctttggatg ctggtgctgc tgaaccatt 60

- 34 -

gccagccggg cctccaggct gaatctgttc ttccagggga aaccaccctt tatgactcaa 120
 cagcagatgt ctctcttttc ccgagaaggg atattagatg ccctctttgt tctctttgaa 180
 gaatgcagtc agcctgtctt gatgaagatt aagcacgtga gcaactttgt ccggaagtat 240
 tccgacacca tagctgagtt acaggagctc cagccttcgg caaaggactt cgaagtcaga 300
 agtctttag gttgtggta ctttctgtaa gtgcagggtg taagagagaa agcaaccggg 360
 gacatctatg ctatgaaagt gatgaagaag aaggctttat tggcccagga gcaggtttca 420
 ttttttgagg aagagcggaa catattatct cgaagcacia gcccgtagg cccccaatta 480
 cagtatgcct ttcaggacaa aaatcacctt tatctggta tggaatatca gcctggaggg 540
 gacttgctgt cacttttgaa tagatatgag gaccagttag atgaaaacct gatacagttt 600
 tacctagctg agctgatttt ggctgttcac agcgttcac tgatgggata cgtgcacga 660
 gacatcaagc ctgagaacat tctcgttgac cgcacaggac acatcaagct ggtggatttt 720
 ggatctgccg cgaaaatgaa ttcaaacaag atggtgaatg ccaaactccc gattgggacc 780
 ccagattaca tggctcctga agtctgact gtgatgaacg gggatggaaa aggcacctac 840
 ggctggact gtgactggtg gtcagtgggc gtgattgcct atgagatgat ttatgggaga 900
 tcccccttcg cagaggggaa ctctgccaga accttcaata acattatgaa tttccagcgg 960
 tttttgaaat ttccagatga ccccaaagtg agcagtgact ttcttgatct gattcaaagc 1020
 ttgttggtgcg gccagaaaga gagaactgaag tttgaaggta tttgtgcca tctttcttc 1080
 tctaaaattg actggaacaa cattcgtaac tctctctccc ccttcgttcc caccctcaag 1140
 tctgacgatg acacotccaa tttgatgaa ccagagaaga attcgtgggt ttcctctct 1200
 ccgtgccagc tgagcccctc aggettotcg ggtgaagaac tgccgtttgt ggggttttcg 1260

- 35 -

tacagcaagg cactggggat tcttggtaga tctgagtctg ttgtgtcggg tctggactcc 1320

octgccaaga ctagctccat ggaaaagaaa cttctcatca aaagcaaaga gctacaagac 1380

tctcaggaca agtgtcacao gatggagcag gaaatgaccc ggttacatcg gagagtgtca 1440

gaggtggagg ctgtgcttag tcagaaggag gtggagctga aggcctctga gactcagaga 1500

tccctcctgg agcaggacct tgctacctac atcacagaat gcagtagctt aaagcgaagt 1560

ttggagcaag cacggatgga ggtgtcccag gaggatgaca aagcactgca gcttctccat 1620

gatatcagag agcagagccg gaagctcaa gaaatcaaag agcaggagta ccaggctcaa 1680

gtggaagaaa tgaggttgat gatgaatcag ttggaagagg atottgtctc agcaagaaga 1740

cggagtgatc tctacgaatc tgagctgaga gagtctcggc ttgctgctga agaattcaag 1800

cggaaagcga cagaatgtca gcataaactg ttgaaggcta aggatcaagg gaagcctgaa 1860

gtgggagaat atgogaaact ggagaagatc aatgctgagc agcagctcaa aattcaggag 1920

ctccaagaga aactggagaa ggctgtaaaa gccagcacgg aggccacoga gctgctgcag 1980

aatatccgcc aggcaaagga gcgagccgag agggagctgg agaagctgca gaaccgagag 2040

gattcttctg aaggcatcag aaagaagctg gtggaagctg aggaacgccg ccattctctg 2100

gagaacaagg taaagagact agagaccatg gagcgtagag aaaacagact gaaggatgac 2160

atccagacaa aatcccaaca gatccagcag atggctgata aaattctgga gctcgaagag 2220

aaacatcggg aggcccaagt ctgagcccag cacctagaag tgcacctgaa acagaaagag 2280

cagcactatg aggaaaagat taaagtgttg gacaatcaga taaagaaaga cctggctgac 2340

aaggagacac tggagaacat gatgcagaga cacgaggagg aggccatga gaagggcaaa 2400

attctcagcg aacagaaggc gatgatcaat gctatggatt ccaagatcag atccctggaa 2460

- 36 -

cagaggattg tggaaactgtc tgaagccaat aaacttgcag caaatagcag tttttttacc 2520

caaaggaaca tgaaggccca agaagagatg atttctgaac tcaggcaaca gaaattttac 2580

ctggagacac aggctgggaa gttggaggcc cagaaccgaa aactggagga gcagctggag 2640

aagatcagcc accaagacca cagtgacaag aatcggtgc tggaaactgga gacaagattg 2700

cgggaggtca gtctagagca cgaggagcag aaactggagc tcaagcgcca gctcacagag 2760

ctacagctct ccctgcagga gcgcgagtca cagttgacag ccctgcaggc tgcacgggcg 2820

gccctggaga gccagcttog ccaggcgaag acagagctgg aagagaccac agcagaagct 2880

gaagaggaga tccaggcaact cacggcacat agagatgaaa tccagcgcaa atttgatgct 2940

cttogtaaca gctgtactgt aatcacagac ctggaggagc agctaaacca gctgaccgag 3000

gacaacgctg aactcaacaa ccaaaacttc tacttgtcca aacaactcga tgaggcttct 3060

ggcgccaacg acgagattgt acaactgcga agtgaagtgg accatctccg ccgggagatc 3120

acggaacgag agatgcagct taccagccag aagcaaacga tggaggctct gaagaccacg 3180

tgcaccatgc tggaggaaca ggtcatggat ttggaggccc taaacgatga gctgctagaa 3240

aaagagcggc agtgggaggc ctggaggagc gtcctgggtg atgagaaatc ccagtttgag 3300

tgctgggttc gagagctgca gagaatgctg gacaccgaga aacagagcag ggcgagagcc 3360

gatcagcgga tcaccagatc tcgccagggtg gtggagctgg cagtgaagga gcacaaggct 3420

gagattctcg ctctgcagca ggctctcaaa gagcagaagc tgaaggccga gagcctctct 3480

gacaagctca atgacctgga gaagaagcat gctatgcttg aaatgaatgc ccgaagctta 3540

cagcagaagc tggagaotga acgagagctc aaacagaggc ttctggaaga gcaagccaaa 3600

ttacagcagc agatggacct gcagaaaaat cacattttcc gtctgactca aggactgcaa 3660

- 37 -

gaagctctag atcgggctga totaotgaag acagaaagaa gtgacttgga gtatcagctg 3720

gaaaacatto aggttctcta ttctoatgaa aaggtgaaaa tggaaggcac tattttotcaa 3780

caaaccaaac tcattgattt totgcaagcc aaaatggacc aacctgctaa aaagaaaaag 3840

gttcctotgc agtacaatga gctgaagctg gccctggaga aggagaaagc tcgctgtgca 3900

gagctagagg aagcccttca gaagaccgc atcgagctcc ggtccgcccg ggaggaagct 3960

gccaccgca aagcaacgga ccaccacac ccatccacgc cagccaccgc gaggcagcag 4020

atcgccatgt ccgccatcgt gcggtcgcca gagcaccagc ccagtgccat gagcctgctg 4080

gccccgccat ccagccgcag aaaggagtct tcaactccag aggaatttag tcggcgtctt 4140

aaggaaagca tgcaccacaa tattoctcac cgattcaacg taggactgaa catgcgagcc 4200

acaaagtgtg ctgtgtgtct ggataccgtg cactttggac gccaggcatc caaatgtctc 4260

gaatgtcagg tgatgtgtca cccaagtgc tccacgtgct tgccagccac ctgcggcttg 4320

cctgtgtaat atgccacaca ottcaccgag gcctttctgcc gtgacaaaat gaactcccca 4380

ggtctccaga ccaaggagcc cagcagcagc ttgcacctgg aagggtggat gaaggtgccc 4440

aggaataaca aacgaggaca gcaaggctgg gacaggaagt acattgtcct ggagggatca 4500

aaagtcctca tttatgacaa tgaagccaga gaagctggac agaggccggt ggaagaattt 4560

gagctgtgcc ttcccgacgg ggatgtatct attcatggtg ccgttggtgc ttccgaactc 4620

gcaaatacag ccaaagcaga tgtcccatc atactgaaga tggaatctca cccgcacacc 4680

acctgctggc ccgggagaac cctotacttg ctagctccca gcttcctga caaacagcgc 4740

tggtgcaccg ccttagaatc agttgtcgca ggtgggagag tttctaggga aaaagcagaa 4800

gctgatgcta aactgcttgg aaactccctg ctgaaactgg aaggtgatga ccgtctagac 4860

- 38 -

atgaactgca cgctgccott cagtgaccag gtggtgttgg tgggcaccga ggaagggctc 4920
tacgccctga atgtcttgaa aaactcccta acccatgtcc caggaattgg agcagtotto 4980
caaatttata ttatcaagga cctggagaag ctactcatga tagcaggaga agagcgggca 5040
ctgtgtcttg tggacgtgaa gaaagtgaaa cagtccctgg cccagtccca cctgcoctgcc 5100
cagcccgaca tctcacccaa cttttttgaa gctgtcaagg gctgccactt gtttggggca 5160
ggcaagattg agaacgggct ctgcatctgt gcagccatgc ccagcaaagt cgtcattctc 5220
cgctacaacg aaaacctcag caaatactgc atcoggaaag agatagagac ctcagagccc 5280
tgcagctgta tccacttcac caattacagt atctcattg gaaccaataa attctacgaa 5340
atcgacatga agcagtacac gctcgaggaa ttcttgata agaatgacca ttcttggca 5400
cctgctgtgt ttgccgctc ttccaacagc ttccctgtct caatcgtgca ggtgaacagc 5460
gcagggcagc gagaggagta cttgctgtgt ttccacgaat ttggagtgtt cgtggattct 5520
tacggaagac gtagccgcac agacgatctc aagtggagtc gttaccttt ggcctttgcc 5580
tacagagaac cctatctgtt tgtgaccac ttcaactcac tcgaagtaat tgagatccag 5640
gcacgctcct cagcaggagc cctgcccga ggtacctgg acatcccga cccgcgctac 5700
ctgggccctg ccatttcctc aggagcgatt tacttggcgt cctcatacca ggataaatta 5760
agggtcattt gctgcaaggg aaacctcgtg aaggagtccg gcaactgaaca ccaccggggc 5820
ccgtccacct cccgcagcag cccaacaag cgaggccac ccacgtacaa cgagcacatc 5880
accaagcgcg tggcctccag cccagcgccg ccgaaggcc ccagccaccc gcgagagcca 5940
agcacacccc accgctaccg cgaggggcgg accgagctgc gcagggacaa gtotcctggc 6000
cgccccctgg agcgagagaa gtccccggc cggatgctca gcacgcggag agagcggctc 6060

- 39 -

cccgggaggc tgtttgaaga cagcagcagg ggcgggctgc ctgcgggagc cgtgaggacc 6120

ccgctgtccc aggtgaacaa ggtctgggac cagtcttcag tataaatctc agccagaaaa 6180

accaactcct catcttgatc tgcaggaaaa caccaaacac actatggaac tctgctgatg 6240

gggacccaag cgcacacgtg ctcagccacc ctctggctca gcggggccca gaccacctc 6300

ggcacggaca cccctgtctc caggaggggc aggtggctga ggctcttcgg agctgtcagc 6360

gcccggtgcc tgcctgggac acctccctgc agtcctctct ttgcaotttg ttactcttct 6420

aaagcattca caaacttttg tacctagctc tagcctgtac cagttagtct atcaaaggaa 6480

accaaccggg atgctaacaa caacatgggt agaatcctaa ttagctactt taagatccta 6540

ggattgggtg gtttttcttt ttttttctc tttgtttctt tccttttttt tttttttttt 6600

taagacaaca gaattottaa tagatttgaa tagcgacgta ttctctgttg tagtcatttt 6660

tagctcgacc acatcatcag gtctttgccà ccgaggcata gtgtagaaca gtcccggta 6720

gttggccaaac ctcccgcagc caagtaggtt catccttggt cctgttcatt ctcatagatg 6780

gcctgtcttt ccccgagggtg acatcgtagc caaatgttta ctgttttcat tgccttttat 6840

ggccttgacg acttccctc ccaccagctg agaatgtatg gaggtcatcg gggcctcagc 6900

tggaggcag tgaactgggg ccaagggacc tcgagacgct ttccttcccc acccccagc 6960

gtcatctccc cagcctgctg ttcccgttt ccatatagct ttggccagga aagcatgcaa 7020

tagaactgct cggagccag cactcctggg tctcggggc ggggagggga cgggggcacc 7080

cacttccttg tctgtgacgg cgtgttgctt cccactctg gatggggaag aggcccgctg 7140

ggagtctctg atggcagttc actgcatgtg ctgccccctt gggttgctct gccaatgtat 7200

taataccatc ccatagctcc tgccaaatcg agaccctctg acgaactgac gactaactgg 7260

- 40 -

ccaccacaag ctgcagtctg tagcactgaa .caaacaaaaa acaaaacgct caagccttac 7320

gaccagagaa ggatttcagc aaaccaccac ctcccactca gtgtcccctc caaacttcac 7380

acttccctgc ctgcagagga tgactctgtt cacaccaat ccagcgcggg tctacccac 7440

gaaactgtga ctttccaaat gagcctttcc ctagggctag acctaagacc aggaagtttg 7500

agaaagcagc cgcagctcaa ctcttcagc tccgccaggg ttgggaagtc cttaggtgca 7560

gtgcggctcc cactgggtot gcgaccctc ctattagagt acgaaattcc tggcaactgg 7620

tatagaacca acctagaggo tttgcagttg gcaagctaac tcggggcctt atttctgcct 7680

ttaatctccc acaaggcctc tgttgctttg ggtcctccac gactottagg cccgcctcaa 7740

caaccaggc acctcctagg taggctcaaa ggtagaccg tttccaccgc agcaggtgaa 7800

catgaccgtg ttttcaactg tgtccacagt tcagatccct ttccagattg caacctggcc 7860

tgcatcccag ctccctcctg ctctgtctt aacctaatg ctttcttggt tgaaacgcct 7920

acaaacctcc atgtggtagc tcctttggca aatgtcctgc tgtggcggtt tatgtgttgc 7980

ttggagtctg tggggtogta ctccctcccc tcccgctccc agggcagatt tgattgaatg 8040

tttgctgaag ttttgtctct tggccacag tatttgaaa ggtcactgaa aatgggtctt 8100

tcagtcttgg catttcattt aggatctcca tgagaaatgg gcttcttgag cctgaaaaat 8160

gtatattgtg tgtctcatct gtgaactgct ttctgctata tagaactagc taaaagact 8220

gtacatattt acaagaaact ttatatctgt aaaaaaaaaa agaggaaatt gaattgggtt 8280

ctactttttt attgtaaaag gtgcattttt caacacttac ttttggtttc aatgggtggt 8340

gttggtggaca gccatcttca ctggagggtg gggagctccg tgtgaccacc aagatgccag 8400

caggatatac cgtaacacga aattgctgtc aaaagcttat tagcatcaat caagattota 8460

- 41 -

gggtotccaaa agtacagget ttttcttcat tacctttttt attcagaacg aggaagagaa 8520
 cacaaggaat gattcaagat ccaccttgag aggaatgaac tttgttggtg aacaattagt 8580
 gaaataaagc aatgatctaa act 8603

<210> 5
 <211> 1286
 <212> PRT
 <213> Homo sapiens

<400> 5

Val Leu Asp Asn Gln Ile Lys Lys Asp Leu Ala Asp Lys Glu Thr Leu
 1 5 10 15

Glu Asn Met Met Gln Arg His Glu Glu Glu Ala His Glu Lys Gly Lys
 20 25 30

Ile Leu Ser Glu Gln Lys Ala Met Ile Asn Ala Met Asp Ser Lys Ile
 35 40 45

Arg Ser Leu Glu Gln Arg Ile Val Glu Leu Ser Glu Ala Asn Lys Leu
 50 55 60

Ala Ala Asn Ser Ser Leu Phe Thr Gln Arg Asn Met Lys Ala Gln Glu
 65 70 75 80

Glu Met Ile Ser Glu Leu Arg Gln Gln Lys Phe Tyr Leu Glu Thr Gln
 85 90 95

Ala Gly Lys Leu Glu Ala Gln Asn Arg Lys Leu Glu Glu Gln Leu Glu
 100 105 110

- 42 -

Lys Ile Ser His Gln Asp His Ser Asp Lys Asn Arg Leu Leu Glu Leu
115 120 125

Glu Thr Arg Leu Arg Glu Val Ser Leu Glu His Glu Glu Gln Lys Leu
130 135 140

Glu Leu Lys Arg Gln Leu Thr Glu Leu Gln Leu Ser Leu Gln Glu Arg
145 150 155 160

Glu Ser Gln Leu Thr Ala Leu Gln Ala Ala Arg Ala Ala Leu Glu Ser
165 170 175

Gln Leu Arg Gln Ala Lys Thr Glu Leu Glu Glu Thr Thr Ala Glu Ala
180 185 190

Glu Glu Glu Ile Gln Ala Leu Thr Ala His Arg Asp Glu Ile Gln Arg
195 200 205

Lys Phe Asp Ala Leu Arg Asn Ser Cys Thr Val Ile Thr Asp Leu Glu
210 215 220

Glu Gln Leu Asn Gln Leu Thr Glu Asp Asn Ala Glu Leu Asn Asn Gln
225 230 235 240

Asn Phe Tyr Leu Ser Lys Gln Leu Asp Glu Ala Ser Gly Ala Asn Asp
245 250 255

Glu Ile Val Gln Leu Arg Ser Glu Val Asp His Leu Arg Arg Glu Ile
260 265 270

- 43 -

Thr Glu Arg Glu Met Gln Leu Thr Ser Gln Lys Gln Thr Met Glu Ala
275 280 285

Leu Lys Thr Thr Cys Thr Met Leu Glu Glu Gln Val Met Asp Leu Glu
290 295 300

Ala Leu Asn Asp Glu Leu Leu Glu Lys Glu Arg Gln Trp Glu Ala Trp
305 310 315 320

Arg Ser Val Leu Gly Asp Glu Lys Ser Gln Phe Glu Cys Arg Val Arg
325 330 335

Glu Leu Gln Arg Met Leu Asp Thr Glu Lys Gln Ser Arg Ala Arg Ala
340 345 350

Asp Gln Arg Ile Thr Glu Ser Arg Gln Val Val Glu Leu Ala Val Lys
355 360 365

Glu His Lys Ala Glu Ile Leu Ala Leu Gln Gln Ala Leu Lys Glu Gln
370 375 380

Lys Leu Lys Ala Glu Ser Leu Ser Asp Lys Leu Asn Asp Leu Glu Lys
385 390 395 400

Lys His Ala Met Leu Glu Met Asn Ala Arg Ser Leu Gln Gln Lys Leu
405 410 415

Glu Thr Glu Arg Glu Leu Lys Gln Arg Leu Leu Glu Glu Gln Ala Lys
420 425 430

- 44 -

Leu Gln Gln Gln Met Asp Leu Gln Lys Asn His Ile Phe Arg Leu Thr
435 440 445

Gln Gly Leu Gln Glu Ala Leu Asp Arg Ala Asp Leu Leu Lys Thr Glu
450 455 460

Arg Ser Asp Leu Glu Tyr Gln Leu Glu Asn Ile Gln Val Leu Tyr Ser
465 470 475 480

His Glu Lys Val Lys Met Glu Gly Thr Ile Ser Gln Gln Thr Lys Leu
485 490 495

Ile Asp Phe Leu Gln Ala Lys Met Asp Gln Pro Ala Lys Lys Lys Lys
500 505 510

Val Pro Leu Gln Tyr Asn Glu Leu Lys Leu Ala Leu Glu Lys Glu Lys
515 520 525

Ala Arg Cys Ala Glu Leu Glu Glu Ala Leu Gln Lys Thr Arg Ile Glu
530 535 540

Leu Arg Ser Ala Arg Glu Glu Ala Ala His Arg Lys Ala Thr Asp His
545 550 555 560

Pro His Pro Ser Thr Pro Ala Thr Ala Arg Gln Gln Ile Ala Met Ser
565 570 575

Ala Ile Val Arg Ser Pro Glu His Gln Pro Ser Ala Met Ser Leu Leu
580 585 590

- 45 -

Ala Pro Pro Ser Ser Arg Arg Lys Glu Ser Ser Thr Pro Glu Glu Phe
595 600 605

Ser Arg Arg Leu Lys Glu Arg Met His His Asn Ile Pro His Arg Phe
610 615 620

Asn Val Gly Leu Asn Met Arg Ala Thr Lys Cys Ala Val Cys Leu Asp
625 630 635 640

Thr Val His Phe Gly Arg Gln Ala Ser Lys Cys Leu Glu Cys Gln Val
645 650 655

Met Cys His Pro Lys Cys Ser Thr Cys Leu Pro Ala Thr Cys Gly Leu
660 665 670

Pro Ala Glu Tyr Ala Thr His Phe Thr Glu Ala Phe Cys Arg Asp Lys
675 680 685

Met Asn Ser Pro Gly Leu Gln Thr Lys Glu Pro Ser Ser Ser Leu His
690 695 700

Leu Glu Gly Trp Met Lys Val Pro Arg Asn Asn Lys Arg Gly Gln Gln
705 710 715 720

Gly Trp Asp Arg Lys Tyr Ile Val Leu Glu Gly Ser Lys Val Leu Ile
725 730 735

Tyr Asp Asn Glu Ala Arg Glu Ala Gly Gln Arg Pro Val Glu Glu Phe
740 745 750

- 46 -

Glu Leu Cys Leu Pro Asp Gly Asp Val Ser Ile His Gly Ala Val Gly
755 760 765

Ala Ser Glu Leu Ala Asn Thr Ala Lys Ala Asp Val Pro Tyr Ile Leu
770 775 780

Lys Met Glu Ser His Pro His Thr Thr Cys Trp Pro Gly Arg Thr Leu
785 790 795 800

Tyr Leu Leu Ala Pro Ser Phe Pro Asp Lys Gln Arg Trp Val Thr Ala
805 810 815

Leu Glu Ser Val Val Ala Gly Gly Arg Val Ser Arg Glu Lys Ala Glu
820 825 830

Ala Asp Ala Lys Leu Leu Gly Asn Ser Leu Leu Lys Leu Glu Gly Asp
835 840 845

Asp Arg Leu Asp Met Asn Cys Thr Leu Pro Phe Ser Asp Gln Val Val
850 855 860

Leu Val Gly Thr Glu Glu Gly Leu Tyr Ala Leu Asn Val Leu Lys Asn
865 870 875 880

Ser Leu Thr His Val Pro Gly Ile Gly Ala Val Phe Gln Ile Tyr Ile
885 890 895

Ile Lys Asp Leu Glu Lys Leu Leu Met Ile Ala Gly Glu Glu Arg Ala
900 905 910

- 47 -

Leu Cys Leu Val Asp Val Lys Lys Val Lys Gln Ser Leu Ala Gln Ser
915 920 925

His Leu Pro Ala Gln Pro Asp Ile Ser Pro Asn Ile Phe Glu Ala Val
930 935 940

Lys Gly Cys His Leu Phe Gly Ala Gly Lys Ile Glu Asn Gly Leu Cys
945 950 955 960

Ile Cys Ala Ala Met Pro Ser Lys Val Val Ile Leu Arg Tyr Asn Glu
965 970 975

Asn Leu Ser Lys Tyr Cys Ile Arg Lys Glu Ile Glu Thr Ser Glu Pro
980 985 990

Cys Ser Cys Ile His Phe Thr Asn Tyr Ser Ile Leu Ile Gly Thr Asn
995 1000 1005

Lys Phe Tyr Glu Ile Asp Met Lys Gln Tyr Thr Leu Glu Glu Phe
1010 1015 1020

Leu Asp Lys Asn Asp His Ser Leu Ala Pro Ala Val Phe Ala Ala
1025 1030 1035

Ser Ser Asn Ser Phe Pro Val Ser Ile Val Gln Val Asn Ser Ala
1040 1045 1050

Gly Gln Arg Glu Glu Tyr Leu Leu Cys Phe His Glu Phe Gly Val
1055 1060 1065

- 48 -

Phe Val Asp Ser Tyr Gly Arg Arg Ser Arg Thr Asp Asp Leu Lys
1070 1075 1080

Trp Ser Arg Leu Pro Leu Ala Phe Ala Tyr Arg Glu Pro Tyr Leu
1085 1090 1095

Phe Val Thr His Phe Asn Ser Leu Glu Val Ile Glu Ile Gln Ala
1100 1105 1110

Arg Ser Ser Ala Gly Thr Pro Ala Arg Ala Tyr Leu Asp Ile Pro
1115 1120 1125

Asn Pro Arg Tyr Leu Gly Pro Ala Ile Ser Ser Gly Ala Ile Tyr
1130 1135 1140

Leu Ala Ser Ser Tyr Gln Asp Lys Leu Arg Val Ile Cys Cys Lys
1145 1150 1155

Gly Asn Leu Val Lys Glu Ser Gly Thr Glu His His Arg Gly Pro
1160 1165 1170

Ser Thr Ser Arg Ser Ser Pro Asn Lys Arg Gly Pro Pro Thr Tyr
1175 1180 1185

Asn Glu His Ile Thr Lys Arg Val Ala Ser Ser Pro Ala Pro Pro
1190 1195 1200

Glu Gly Pro Ser His Pro Arg Glu Pro Ser Thr Pro His Arg Tyr
1205 1210 1215

- 49 -

Arg Glu Gly Arg Thr Glu Leu Arg Arg Asp Lys Ser Pro Gly Arg
 1220 1225 1230

Pro Leu Glu Arg Glu Lys Ser Pro Gly Arg Met Leu Ser Thr Arg
 1235 1240 1245

Arg Glu Arg Ser Pro Gly Arg Leu Phe Glu Asp Ser Ser Arg Gly
 1250 1255 1260

Arg Leu Pro Ala Gly Ala Val Arg Thr Pro Leu Ser Gln Val Asn
 1265 1270 1275

Lys Val Trp Asp Gln Ser Ser Val
 1280 1285

<210> 6

<211> 5261

<212> DNA

<213> Homo sapiens

<400> 6

cagagcaggg cgagagccga tcagcggatc accgagtctc gccaggtggt ggagctggca 60

gtgaaggagc acaaggctga gattctcgct ctgcagcagg ctctcaaaga gcagaagctg 120

aaggccgaga gcctctctga caagctcaat gacctggaga agaagcatgc tatgcttgaa 180

atgaatgccc gaagottaca gcagaagctg gagactgaac gagagctcaa acagaggctt 240

ctggaagagc aagccaaatt acagcagcag atggacctgc agaaaaatca cattttccgt 300

ctgactcaag gactgcaaga agctctagat cgggctgata tactgaagac agaaagaagt 360

- 50 -

gaattggagt atcagctgga aaacattcag gttctctatt ctcatgaaaa ggtgaaaatg 420

gaaggcacta tttctcaaca aoccaaactc attgattttc tgcaagccaa aatggaccaa 480

cctgctaaaa agaaaaaggt tcctctgcag tacaatgagc tgaagctggc cctggagaag 540

gagaaagctc gctgtgcaga gctagaggaa gcccttcaga agaccgcgat cgagctccgg 600

tcggcccggg aggaagctgc ccacgcgcaa gcaacggacc acccacacco atccacgcca 660

gccacgcga ggcagcagat cgccatgtct gccatcgtgc ggtcgccaga gcaccagccc 720

agtgccatga gcctgctggc ccgcgccatcc agccgcagaa aggagtcttc aactccagag 780

gaatttagtc ggcgtcttaa ggaacgcatt caccacaata ttcctcacgc attcaacgta 840

ggactgaaca tgcgagccac aaagtgtgct gtgtgtctgg atacctgca ctttgacgc 900

caggcatcca aatgtctoga atgtcaggtg atgtgtcacc ccaagtgtc cagtgcttg 960

ccagccacct gcggcttgcc tgctgaatat gccacacact tcaccgaggc cttctgccgt 1020

gacaaaatga actcccagg tctccagacc aaggagccca gcagcagctt gcacctggaa 1080

gggtggatga aggtgcccag gaataacaaa cgaggacagc aaggctggga caggaaagta 1140

attgtcctgg agggatcaaa agtcctcatt tatgacaatg aagccagaga agctggacag 1200

aggccggtgg aagaatttga gctgtgcctt ccgcacgggg atgtatctat tcatggtgcc 1260

gttggtgctt ccgaactcgc aaatacagcc aaagcagatg tccatacat actgaagatg 1320

gaatctcacc cgcacaccac ctgctggccc gggagaaccc totacttgct agctcccagc 1380

ttccctgaca aacagcgctg ggtcacgcgc ttagaatcag ttgtcgcagg tgggagagtt 1440

tctagggaaa aagcagaagc tgatgctaaa ctgcttgga actccctgct gaaactggaa 1500

ggtgatgacc gtctagacat gaactgcacg ctgcccttca gtgaccaggt ggtgttggtg 1560

- 51 -

ggcaccgagg aagggctcta cgccctgaat gtcttgaaaa actccctaac ccatgtccca 1620

ggaattggag cagtcttcca aatttatatt atcaaggacc tggagaagct actcatgata 1680

gcaggagaag agcgggcact gtgtcttggt gacgtgaaga aagtgaacaa gtccctggcc 1740

cagtcccacc tgcctgcca gcccgacatc tcaccaaca tttttgaagc tgtcaagggc 1800

tgccacttgt ttggggcagg caagattgag aacgggctct gcctctgtgc agccatgcc 1860

agcaaagtgc tcattctcgc ctacaacgaa aacctcagca aatactgcat ccggaaagag 1920

atagagacct cagagccctg cagctgtatc cacttcacca attacagtat cctcattgga 1980

accaataaat tctacgaaat cgacatgaag cagtacacgc tcgaggaatt cctggataag 2040

aatgaccatt ccttggcacc tgctgtgttt gccgcctctt ccaacagctt ccctgtctca 2100

atcgtgcagg tgaacagcgc agggcagcga gaggagtact tgctgtgttt ccacgaattt 2160

ggagtgttcg tggattctta cggaagacgt agccgcacag acgatctcaa gtggagtgcg 2220

ttacotttgg cctttgccta cagagaacct tatctgtttg tgaccactt caactcactc 2280

gaagtaattg agatccaggc acgctcctca gcagggacct ctgcccagac gtacctggac 2340

atcccgaacc cgcgctacct gggccctgcc atttctcag gagcgattta cttggcgctc 2400

tcataccagg ataaattaag ggtcatttgc tgcaaggga acctcgtgaa ggagtccggc 2460

actgaacacc accggggccc gtccacctc cgcagcagcc ccaacaagcg agggccacc 2520

acgtacaacg agcacatcac caagcgcgtg gcctccagcc cagcgccgcc cgaaggcccc 2580

agccaccgcg gagagccaag cacacccac cgctaccgcg aggggcggac cgagctgcgc 2640

agggacaagt ctctggccg cccctggag cgagagaagt ccccgggccg gatgctcagc 2700

acgcggagag agcgtcccc cgggaggtg tttgaagaca gcagcagggg ccggctgcct 2760

- 52 -

gcgggagccg tgaggacccc gctgtcccag gtgaacaagg tctgggacca gtcttcagta 2820

taaattctcag ccagaaaaac caactcctca tottgatctg caggaaaaca ccaaacacac 2880

tatggaactc tgctgatggg gaccaagcg cccacgtgct cagccaccct ctggctcagc 2940

ggggcccaga cccacctcgg cacggacacc cctgtctcca ggaggggcag gtggctgagg 3000

ctcttcggag ctgtcagcgc cgggtgcctg ccctgggcac ctccctgcag tcattctctt 3060

gcactttgtt actotttcaa agcattcaca aacttttgta cctagctcta gcctgtacca 3120

gttagttcat caaaggaaac caaccgggat gctaacaaca acatggtttag aatcctaatt 3180

agctacttta agatcctagg attggttggt tttctttttt tttttctott tgtttcttto 3240

cttttttttt ttttttttta agacaacaga attcttaata gatttgaata gcgacgtatt 3300

tcctgttgta gtcattttta gctcgaccac atcatcaggt ctttgccacc gaggcatagt 3360

gtagaacagt cccggtcagt tggccaacct cccgcagcca agtaggttca tccttgttcc 3420

tgttcattct catagatggc cctgctttcc ccagggtgac atcgtagcca aatgtttact 3480

gttttcattg ccttttatgg ccttgacgac ttccctccc accagctgag aatgtatgga 3540

ggtcacggg gcctcagctc ggaggcagtg acttggggcc aagggaacct gagacgcttt 3600

ccttccccac ccccagcgt catctcccca gcctgctgtt cccgctttcc atatagcttt 3660

ggccaggaaa gcatgcaata gacttgctcg gagcccagca ctctgggto tcggggtcgg 3720

ggaggggacg ggggcaccca cttccttgto tgtgacggcg tgttgttccc cactctggga 3780

tggggaagag gcccgtcggg agttctgcat ggcagttcac tgcattgtct gccccttg 3840

gttgctctgo caatgtatta ataccatccc atagctcctg ccaaactgag accctctgac 3900

gacttgccga ctaactggcc accacaagct gcagtctgta gcaactgaaca aacaaaaaac 3960

- 53 -

aaaacgctca agccttacga ccagagaagg atttcagcaa accaccacct cccactcagt 4020
gtccctcca aacttcacac ttccctgcoct gcagaggatg actctgttca cacccaatcc 4080
agcgcggttc taccacacga aactgtgact ttccaaatga gcctttccct agggctagac 4140
ctaagaccag gaagtttgag aaagcagccg cagotcaact ottccagctc cgccagggtt 4200
gggaagtcoct taggtgcagt gcggtccca ctgggtctgc ggaccctcct attagagtac 4260
gaaattcoctg gcaactggta tagaaccaac ctagaggctt tgcagttggc aagctaacto 4320
goggccttat ttctgccttt aatctccac aaggcatctg ttgctttggg tcctccacga 4380
ctcttaggcc cgcctcaaca acccaggcac ctctaggta ggctcaaagg tagaccogtt 4440
tccaccgcag caggtgaaca tgaccgtgtt ttcaactgtg tccacagttc agatcccttt 4500
ccagattgca acctggcctg catcccagct ccttcctgct cgtgtcttaa cctaagtgt 4560
ttcttgtttg aaacgcctac aaacotocat gtggtagctc ctttggtcaa tgtcctgctg 4620
tggcgtttta tgtgttgctt ggagtctgtg gggctgtaact cctccctc cgtccccag 4680
ggcagatttg attgaatgtt tgctgaagtt ttgtctcttg gtccacagta tttggaaagg 4740
tcactgaaaa tgggtctttc agtcttgga tttcatttag gatctccatg agaaatgggc 4800
ttcttgagcc ctgaaaatgt atattgtgtg tctcatctgt gaactgcttt ctgctatata 4860
gaaactagctc aaaagactgt acatatttac aagaaacttt atattcgtaa aaaaaaaaaag 4920
aggaaattga attggtttct acttttttat tgtaaaaggc gcatttttca acaattactt 4980
ttggtttcaa tgggtgtagt tgtggacagc catcttcaact ggaggggtggg gagctccgtg 5040
tgaccaccaa gatgccagca ggatataccg taacacgaaa ttgctgtcaa aagcttatta 5100
gcatcaatca agattctagg totccaaaag tacaggcttt ttcttcatta ccttttttat 5160

- 54 -

tcagaacgag gaagagaaca caaggaatga ttcaagatcc acottgagag gaatgaactt 5220

tgttgttgaa caattagtga aataaagcaa tgatctaaac t 5261

<210> 7

<211> 631

<212> PRT

<213> Homo sapiens

<400> 7

Met Ser Ala Glu Val Arg Leu Arg Gln Leu Gln Gln Leu Val Leu Asp
1 5 10 15

Pro Gly Phe Leu Gly Leu Glu Pro Leu Leu Asp Leu Leu Leu Gly Val
20 25 30

His Gln Glu Leu Gly Ala Ser His Leu Ala Gln Asp Lys Tyr Val Ala
35 40 45

Asp Phe Leu Gln Trp Val Glu Pro Ile Ala Ala Arg Leu Lys Glu Val
50 55 60

Arg Leu Gln Arg Asp Asp Phe Glu Ile Leu Lys Val Ile Gly Arg Gly
65 70 75 80

Ala Phe Ser Glu Val Ala Val Val Lys Met Lys Gln Thr Gly Gln Val
85 90 95

Tyr Ala Met Lys Ile Met Asn Lys Trp Asp Met Leu Lys Arg Gly Glu
100 105 110

- 55 -

Val Ser Cys Phe Arg Glu Glu Arg Asp Val Leu Val Lys Gly Asp Arg
115 120 125

Arg Trp Ile Thr Gln Leu His Phe Ala Phe Gln Asp Glu Asn Tyr Leu
130 135 140

Tyr Leu Val Met Glu Tyr Tyr Val Gly Gly Asp Leu Leu Thr Leu Leu
145 150 155 160

Ser Lys Phe Gly Glu Arg Ile Pro Ala Glu Met Ala Arg Phe Tyr Leu
165 170 175

Ala Glu Ile Val Met Ala Ile Asp Ser Val His Arg Leu Gly Tyr Val
180 185 190

His Arg Asp Ile Lys Pro Asp Asn Ile Leu Leu Asp Arg Cys Gly His
195 200 205

Ile Arg Leu Ala Asp Phe Gly Ser Cys Leu Lys Leu Gln Pro Asp Gly
210 215 220

Met Val Arg Ser Leu Val Ala Val Gly Thr Pro Asp Tyr Leu Ser Pro
225 230 235 240

Glu Ile Leu Gln Ala Val Gly Gly Gly Pro Gly Ala Gly Ser Tyr Gly
245 250 255

Pro Glu Cys Asp Trp Trp Ala Leu Gly Val Phe Ala Tyr Glu Met Phe
260 265 270

- 56 -

Tyr Gly Gln Thr Pro Phe Tyr Ala Asp Ser Thr Ala Glu Thr Tyr Ala
275 280 285

Lys Ile Val His Tyr Arg Glu His Leu Ser Leu Pro Leu Ala Asp Thr
290 295 300

Val Val Pro Glu Glu Ala Gln Asp Leu Ile Arg Gly Leu Leu Cys Pro
305 310 315 320

Ala Glu Ile Arg Leu Gly Arg Gly Gly Ala Gly Asp Phe Gln Lys His
325 330 335

Pro Phe Phe Phe Gly Leu Asp Trp Glu Gly Leu Arg Asp Ser Val Pro
340 345 350

Pro Phe Thr Pro Asp Phe Glu Gly Ala Thr Asp Thr Cys Asn Phe Asp
355 360 365

Val Val Glu Asp Arg Leu Thr Ala Met Val Ser Gly Gly Gly Glu Thr
370 375 380

Leu Ser Asp Met Gln Glu Asp Met Pro Leu Gly Val Arg Leu Pro Phe
385 390 395 400

Val Gly Tyr Ser Tyr Cys Cys Met Ala Phe Arg Asp Asn Gln Val Pro
405 410 415

Asp Pro Thr Pro Met Glu Leu Glu Ala Leu Gln Leu Pro Val Ser Asp
420 425 430

- 57 -

Leu Gln Gly Leu Asp Leu Gln Pro Pro Val Ser Pro Pro Asp Gln Val
435 440 445

Ala Glu Glu Ala Asp Leu Val Ala Val Pro Ala Pro Val Ala Glu Ala
450 455 460

Glu Thr Thr Val Thr Leu Gln Gln Leu Gln Glu Ala Leu Glu Glu Glu
465 470 475 480

Val Leu Thr Arg Gln Ser Leu Ser Arg Glu Leu Glu Ala Ile Arg Thr
485 490 495

Ala Asn Gln Asn Phe Ser Ser Gln Leu Gln Glu Ala Glu Val Arg Asn
500 505 510

Arg Asp Leu Glu Ala His Val Arg Gln Leu Gln Glu Arg Met Glu Met
515 520 525

Leu Gln Ala Pro Gly Ala Ala Ala Ile Thr Gly Val Pro Ser Pro Arg
530 535 540

Ala Thr Asp Pro Pro Ser His Leu Asp Gly Pro Pro Ala Val Ala Val
545 550 555 560

Gly Gln Cys Pro Leu Val Gly Pro Gly Pro Met His Arg Arg His Leu
565 570 575

Leu Leu Pro Ala Arg Ile Pro Arg Pro Gly Leu Ser Glu Ala Arg Cys
580 585 590

- 58 -

Leu Leu Leu Phe Ala Ala Ala Leu Ala Ala Ala Ala Thr Leu Gly Cys
 595 600 605

Thr Gly Leu Val Ala Tyr Thr Gly Gly Leu Thr Pro Val Trp Cys Phe
 610 615 620

Pro Gly Ala Thr Phe Ala Pro
 625 630

<210> 8
 <211> 6156
 <212> DNA
 <213> Homo sapiens

<400> 8

atgttgaagt tcaaatatgg agcgcggaat cctttggatg ctggtgctgc tgaacccatt 60

gccagccggg cctccaggct gaatctgttc ttccagggga aaccaccctt tatgactcaa 120

cagcagatgt ctctctttc ccgagaaggg atattagatg ccctctttgt tctctttgaa 180

gaatgcagtc agcctgctct gatgaagatt aagcacgtga gcaactttgt ccggaagtat 240

tccgacacca tagctgagtt acaggagctc cagccttcgg caaaggactt cgaagtcaga 300

agtctttag gtttgtgtca ctttgcgtgaa gtgcagggtg taagagagaa agcaaccggg 360

gacatctatg ctatgaaagt gatgaagaag aaggctttat tggcccagga gcaggtttca 420

ttttttgagg aagagcggaa catattatot cgaagcacaa gccctgtgat cccccaatta 480

cagtatgcct ttcaggacaa aaatcacctt tatctggtca tggaatatca gcctggaggg 540

gacttgctgt cacttttgaa tagatatgag gaccagttag atgaaaacct gatacagttt 600

tacctagctg agctgatttt ggtgtttcac agcgttcac tgatgggata cgtgcatoga 660

- 59 -

gacatcaagc ctgagaacat tctcgttgac cgcacaggac acatcaagct ggtggatttt 720

ggatctgccg cgaaaatgaa ttcaaacaag atggtgaatg ccaaactccc gattgggacc 780

ccagattaca tggctcctga agtgctgact gtgatgaacg gggatggaaa aggcacctac 840

ggcctggact gtgactggtg gtcagtgggc gtgattgcct atgagatgat ttatgggaga 900

tcccccttcg cagagggaac ctctgccaga aacctcaata acattatgaa ttccagcgg 960

tttttgaaat ttccagatga ccccaaagtg agcagtgact ttcttgatct gattcaaagc 1020

ttgttgctgcg gccagaaaga gagactgaag tttgaaggtc tttgctgcc a tctttcttc 1080

tctaaaattg actggaacaa cattcgtaac tctctcccc ccttcgttcc caccctcaag 1140

tctgacgatg acacotccaa ttttgatgaa ccagagaaga attcgtgggt ttcctctct 1200

ccgtgccagc tgagccctc aggtctctcg ggtgaagaac tgccgtttgt ggggttttcg 1260

tacagcaagg cactggggat tcttggtaga tctgagtctg ttgtgtcggg tctggactcc 1320

cctgccaaaga ctagctccat ggaaaagaaa cttctcatca aaagcaaaga gctacaagac 1380

tctcaggaca agtgtcacia gatggagcag gaaatgacct gggtacatcg gagagtgtca 1440

gaggtggagg ctgtgcttag tcagaaggag gtggagctga aggcctctga gactcagaga 1500

tccctcctgg agcaggacct tgctacctac atcacagaat gcagtagctt aaagcgaagt 1560

ttggagcaag caccgatgga ggtgtcccag gaggatgaca aagcaactgca gcttctccat 1620

gatatcagag agcagagccg gaagctccaa gaaatcaaag agcaggagta ccaggtcaa 1680

gtggaagaaa tgaggttgat gatgaatcag ttggaagagg atcttgtctc agcaagaaga 1740

cggagtgate tctacgaatc tgagctgaga gagtctcggc ttgctgctga agaattcaag 1800

cggaaagcga cagaatgtca gcataaactg ttgaaggcta aggatcaagg gaagcctgaa 1860

- 60 -

gtgggagaat atgcgaaact ggagaagatc aatgctgagc agcagctcaa aattcaggag 1920

otccaagaga aactggagaa ggctgtaaaa gccagcacgg aggccacoga gctgctgcag 1980

aatatccgcc aggcaaagga gcgagccgag agggagctgg agaagctgca gaaccgagag 2040

gattcttctg aaggcatcag aaagaagctg gtggaagctg aggaacgccg ccattctctg 2100

gagaacaagg taaagagact agagaccatg gagcgtagag aaaacagact gaaggatgac 2160

atccagacaa aatcccaaca gatccagcag atggctgata aaattctgga gctcgaagag 2220

aaacatcggg aggcccaagt ctgagcccag cacctagaag tgcacctgaa acagaaagag 2280

cagcactatg aggaaaagat taaagtgttg gacaatcaga taaagaaaga cctggctgac 2340

aaggagacac tggagaacat gatgcagaga cagcaggagg aggcccatga gaagggcaaa 2400

attctcagcg aacagaaggc gatgatcaat gctatggatt ccaagatcag atccctggaa 2460

cagaggattg tggaactgtc tgaagccaat aaacttgacg caaatagcag tctttttacc 2520

caaaggaaca tgaaggocca agaagagatg atttctgaac tcaggcaaca gaaattttac 2580

ctggagacac aggctgggaa gttggaggcc cagaaccgaa aactggagga gcagctggag 2640

aagatcagcc accaagacca cagtgacaag aatcggtgc tggaactgga gacaagattg 2700

cgggaggtca gtctagagca cgaggagcag aaactggagc tcaagcgcca gctcacagag 2760

ctacagctct ccctgcagga gcgcgagtca cagttgacag ccctgcaggc tgcacgggag 2820

gccctggaga gccagcttcg ccaggcgaag acagagctgg aagagaccac agcagaagct 2880

gaagaggaga tccaggcact cacggcacat agagatgaaa tccagcgcaa atttgatgct 2940

cttcgtaaca gctgtactgt aatcacagac ctggaggagc agctaaacca gctgaccgag 3000

gacaacgctg aactcaacaa ccaaaacttc tacttgtcca aacaactcga tgaggcttct 3060

- 61 -

ggcgccaacg acgagattgt acaactgoga agtgaagtgg accatctccg ccgggagatc 3120

acggaacgag agatgcagct taccagccag aagcaaacga tggaggctct gaagaccacg 3180

tgcaccatgc tggaggaaca ggtcatggat ttggaggccc taaacgatga gctgctagaa 3240

aaagagcggc agtgggaggc ctggaggagc gtcctgggtg atgagaaatc ccagtttgag 3300

tgctcgggttc gagagctgca gagaatgctg gacaccgaga aacagagcag ggcgagagcc 3360

gatcagcggg taaccgagtc tcgccagggtg gtggagctgg cagtgaagga gcacaaggct 3420

gagattctcg ctctgcagca ggctctcaaa gagcagaagc tgaaggccga ggcctctct 3480

gacaagctca atgacctgga gaagaagcat gctatgcttg aaatgaatgc ccgaagctta 3540

cagcagaagc tggagactga acgagagctc aaacagaggc ttctggaaga gcaagccaaa 3600

ttacagcagc agatggacct gcagaaaaat cacattttcc gtctgactca aggactgcaa 3660

gaagctctag atcgggctga tctactgaag acagaaagaa gtgacttgga gtatcagctg 3720

gaaaacattc aggttctcta ttctcatgaa aaggtgaaaa tgggaaggcac tatttctcaa 3780

caaaccaaaac tcattgattt tctgcaagcc aaaatggacc aacctgotaa aaagaaaaag 3840

gttctctgac agtacaatga gctgaagctg gccctggaga aggagaaagc tcgctgtgca 3900

gagctagagg aagcccttca gaagaccgc atcgagctcc ggtccgcccg ggaggaagct 3960

gccaccgca aagcaacgga ccaccacac ccattccagc cagccaccgc gaggcagcag 4020

atcgccatgt ccgccatcgt gcggtcgcca gagcaccagc ccagtgccat ggcctgtctg 4080

gccccgccat ccagccgcag aaaggagtct tcaactccag aggaatttag tcggcgtctt 4140

aaggaacgca tgaccacaa tattctcac cgattcaacg taggactgaa catgcgagcc 4200

acaaagtgtg ctgtgtgtct ggataccgtg cactttggac gccaggcatc caaatgtctc 4260

gaatgtcagg tgatgtgtca ccccaagtgc tccacgtgct tgccagccac ctgoggcttg 4320

cctgotgaat atgccacaca cttcaccgag gccttctgcc gtgacaaaat gaactcccca 4380

ggtctccaga ccaaggagcc cagcagcagc ttgcacctgg aagggtggat gaaggtgccc 4440

aggaataaca aacgaggaca gcaaggctgg gacaggaagt acattgtcct ggagggatca 4500

aaagtctca tttatgacaa tgaagccaga gaagctggac agaggccggt ggaagaattt 4560

gagctgtgcc ttcccagcgg ggatgtatct attcatgggtg ccgttgggtgc ttccgaactc 4620

gcaaatacag ccaaagcaga tgtcccatac atactgaaga tggaatctca cccgcacacc 4680

acctgtggc ccgggagaa cctctacttg ctagctccca gcttccctga caaacagcgc 4740

tgggtcaccg cottagaato agttgtcgca ggtgggagag tttctaggga aaaagcagaa 4800

gctgatgcta aactgottgg aaactccctg ctgaaactgg aaggtgatga ccgtctagac 4860

atgaactgca cgtgcccctt cagtgaccag gtggtgttgg tgggcaccga ggaagggctc 4920

tacgccctga atgtcttgaa aaactcccta acccatgtcc caggaattgg agcagtcttc 4980

caaatttata ttatcaagga cctggagaag ctactcatga tagcaggaga agagcgggca 5040

ctgtgtcttg tggacgtgaa gaaagtgaaa cagtccctgg ccagtccca cctgcctgcc 5100

cagcccgaca totcacccaa catttttgaa gctgtcaagg gctgccaatt gtttggggca 5160

ggcaagattg agaacgggct ctgcattgtg gcagccatgc ccagcaaagt cgtcattctc 5220

cgtacaacg aaaacctcag caaatactgc atccggaaag agatagagac ctgagagccc 5280

tgcagctgta tccacttcac caattacagt atcctcattg gaaccaataa attctacgaa 5340

atcgacatga agcagtacac gctcgaggaa ttcctggata agaatgacca ttccttggca 5400

cctgtgtgtg ttgccgccto ttccaacago ttcctgtct caatcgtgca ggtgaacagc 5460

- 63 -

gcagggcagc gagaggagta cttgctgtgt ttccacgaat ttggagtgtt cgtggattct 5520
tacggaagac gtagccgcac agacgatctc aagtggagtc gcttaccttt ggcctttgcc 5580
tacagagaac cctatctgtt tgtgaccac ttcaactcac tcgaagtaat tgagatccag 5640
gcacgtcct cagcagggac ccctgccga gcgtacctgg acatcccgaa cccgcgctac 5700
ctggggccctg ccatttctctc aggagcgatt tacttggcgt cctcatacca ggataaatta 5760
agggtcattt gctgcaaggg aaacctcgtg aaggagtccg gcaactgaaca ccaccggggc 5820
ccgtccacct cccgcagcag cccaacaag cgaggccac ccacgtacaa cgagcacatc 5880
accaagcgcg tggcctccag cccagcgccg ccgaaggcc ccagccacco gcgagagcca 5940
agcacacccc accgtaccg cgagggcgcg accgagctgc gcagggacaa gtctcctggc 6000
cgccccctgg agcgagagaa gtcccccggc cggatgctca gcacgaggag agagcggtcc 6060
ccggggaggc tgtttgaaga cagcagcagg ggccggctgc ctgcgggagc cgtgaggacc 6120
ccgtgtccc aggtgaacaa ggtgaggcag cattcc 6156

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 02/07156

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C07K14/435 C12N15/52 C12N5/10 C12N9/00 C12Q1/68
G01N33/53 G01N33/573 A61P9/10

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C07K C12N C12Q G01N A61P

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, BIOSIS, SEQUENCE SEARCH, PAJ, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 01 38503 A (PLOWMAN GREGORY D ;CLARY DOUGLAS (US); SUGEN INC (US); WHYTE DAVID) 31 May 2001 (2001-05-31) SEQ ID No 1 tables 1,3	1-12, 15-17
X	BARTON G J: "PROTEIN SEQUENCE ALIGNMENT AND DATABASE SCANNING" PROTEIN STRUCTURE PREDICTION. A PRACTICAL APPROACH, XX, XX, 1996, pages 31-63, XP000829540 the whole document	1-12, 15-17
	-/--	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

A document defining the general state of the art which is not considered to be of particular relevance

E earlier document but published on or after the international filing date

L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

O document referring to an oral disclosure, use, exhibition or other means

P document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

G document member of the same patent family

Date of the actual completion of the international search

3 December 2002

Date of mailing of the international search report

11/12/2002

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax (+31-70) 340-3016

Authorized officer

Keller, Y

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 02/07156

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>GEORGE D G ET AL: "CURRENT METHODS IN SEQUENCE COMPARISON AND ANALYSIS" MACROMOLECULAR SEQUENCING AND SYNTHESIS SELECTED METHODS AND APPLICATIONS, XX, XX, 1988, pages 127-149, XP000829541 the whole document</p> <p>----</p>	1-12, 15-17
Y	<p>MADAULE PASCAL ET AL: "A novel partner for the GTP-bound forms of rho and rac" FEBS LETTERS, ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, NL, vol. 377, no. 2, 1995, pages 243-248, XP002200178 ISSN: 0014-5793 the whole document</p> <p>----</p>	1-12, 15-17
Y	<p>DI CUNTO FERDINANDO ET AL: "Citron Rho-interacting kinase, a novel tissue-specific Ser/Thr kinase encompassing the Rho-Rac-binding protein citron" JOURNAL OF BIOLOGICAL CHEMISTRY, AMERICAN SOCIETY OF BIOLOGICAL CHEMISTS, BALTIMORE, MD, US, vol. 273, no. 45, 6 November 1998 (1998-11-06), pages 29706-29711, XP002170360 ISSN: 0021-9258 the whole document</p> <p>----</p>	1-12, 15-17
Y	<p>NAGASE ET AL: "PREDICTION OF THE CODING SEQUENCE OF UNIDENTIFIED HUMAN GENES. XIII. THE COMPLETE SEQUENCE OF 100 NEW CDNA CLONES FROM BRAIN WHICH CODE FOR LARGE PROTEINS IN VITRO" DNA RESEARCH, UNIVERSAL ACADEMY PRESS, JP, vol. 6, 1999, pages 63-70, XP000952912 ISSN: 1340-2838 the whole document</p> <p>-----</p>	1-12, 15-17

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 13, 14 and 15-17 partially

Present claims 13, 14 and 15-17 partially relate to an extremely large number of possible compounds/methods. Support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT is to be found, however, for only a very small proportion of the compounds/products/apparatus/methods claimed. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP 02/07156

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☒ Claims Nos.: 13, 14 and 15-17 partially
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP 02/07156

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0138503	A	31-05-2001	AU
		EP	1926001 A
		WO	1240194 A2
			0138503 A2
			04-06-2001
			18-09-2002
			31-05-2001